



N. ix.

20/2



22200232117







A  
TEXT-BOOK OF PHYSICS



A  
TEXT-BOOK OF PHYSICS

WITH SECTIONS ON THE APPLICATION  
OF PHYSICS TO PHYSIOLOGY  
AND MEDICINE

BY

R. A. LEHFELDT, D.Sc.

PROFESSOR AT THE EAST LONDON TECHNICAL COLLEGE

London  
EDWARD ARNOLD

[ *All rights reserved* ]

[1902]

19 587 045

WELLCOME INSTITUTE LIBRARY	
Coll.	weIMOmec
Call	
No.	Q 400
	1902
	L52t



## PREFACE


THIS book has been written from the point of view of the student of medicine, in the hope of attracting attention to the intimate dependence of physiology on physical principles, and at the same time supplying a practical handbook of moderate dimensions, containing so much of Physics as the student amongst his many other claims can find time for.

The treatment is, however, sufficiently wide to make the book of use to chemical, engineering, and other students, being in fact based on the author's lectures for the Intermediate Examination in Science at London University.

The more difficult and advanced parts have been printed in smaller type, so that they may be omitted if desired.

LONDON,

*March, 1902.*



Digitized by the Internet Archive  
in 2019 with funding from  
Wellcome Library

<https://archive.org/details/b31346613>

# CONTENTS

## CHAPTER I.

### GENERAL PHYSICS.

SECT.	PAGE
1. Conservation of energy . . . . .	1
2. Length and time . . . . .	3
3. Mass . . . . .	8
4. Momentum : force . . . . .	12
5. Work : energy : power . . . . .	20
6. Vectors . . . . .	26
7. Oscillations and waves . . . . .	31
8. Hydrostatics . . . . .	36

## CHAPTER II.

### HEAT.

1. Temperature . . . . .	46
2. Quantity of heat : dynamical equivalent of heat . . . . .	54
3. Thermodynamics . . . . .	62
4. Molecular theory . . . . .	65
5. Conduction of heat . . . . .	68
6. Radiation . . . . .	72

## CHAPTER III.

### PROPERTIES OF FLUIDS.

1. Perfect gases . . . . .	77
2. Condensation and vapour pressure . . . . .	82
3. Liquids . . . . .	88
4. Critical point . . . . .	92
5. Mixtures . . . . .	95
6. Calorimetry of fluids . . . . .	101
7. Surface tension . . . . .	108
8. Flow of liquids . . . . .	113
9. Diffusion : osmosis . . . . .	120

## CHAPTER IV.

### PROPERTIES OF SOLIDS.

1. Elasticity of solids . . . . .	129
2. Limits of elasticity . . . . .	135
3. Thermal properties of solids . . . . .	138
4. Solution . . . . .	142

## CONTENTS

## CHAPTER V.

## SOUND.

SECT.	PAGE
1. Sources of sound . . . . .	145
2. Sound-waves . . . . .	150
3. Resonance . . . . .	157
4. Perception of sound . . . . .	161

## CHAPTER VI.

## CHEMISTRY.

1. Law of mass action . . . . .	164
2. The phase rule . . . . .	170
3. Thermochemistry . . . . .	176
4. Heat and chemical equilibrium . . . . .	180

## CHAPTER VII.

## ELECTRIC CURRENTS.

1. General properties of currents . . . . .	185
2. Electromotive force . . . . .	191
3. Resistance . . . . .	198
4. Distribution of currents . . . . .	206
5. Methods of measurement . . . . .	210
6. Electricity and heat . . . . .	222

## CHAPTER VIII.

## ELECTRO-MAGNETISM.

1. Magnetism . . . . .	226
2. Magnetic field due to a current . . . . .	233
3. Mechanical force on a conductor . . . . .	237
4. Induction of currents . . . . .	239
5. Alternating currents . . . . .	246

## CHAPTER IX.

## ELECTROSTATICS.

1. Forces between electric charges . . . . .	249
2. Electric discharges . . . . .	256

## CHAPTER X.

## LIGHT.

1. Production of light . . . . .	260
2. Light-waves . . . . .	262
3. Phenomena of emission and absorption . . . . .	268
4. Reflection and refraction . . . . .	277
5. Optical instruments . . . . .	285
6. Dispersion . . . . .	291
7. The eye . . . . .	297

INDEX . . . . .	300
-----------------	-----



# A TEXT-BOOK OF PHYSICS

## CHAPTER I.

### GENERAL PHYSICS.

#### § I. Conservation of energy.

THE leading principle running through the science of Physics at the present day is the importance of the part played by the thing known as 'energy' in the phenomena of nature. The conception of energy, derived in the first place from the ability of the muscles to do work, received quantitative definition from a consideration of the work done in simple mechanical instances. When a weight has to be raised the effort involved depends both on the magnitude of the weight and the height through which it has to be raised: if we set the effort as proportional to each of these quantities, the proper measure of it becomes the product of weight into height: this product may be described either as the *work* done or as the *energy* spent; and the justness of this mode of measurement is borne out by other facts. For if the action of some simple machine—say a lever, or a set of pulleys—be studied, it is found easy to arrange that a small weight by falling shall raise a larger one; but that the small weight has to descend through a distance greater in proportion, so that the machine is incapable of giving out more work than is put into it—work being defined as above. Hence though a machine may render the work more convenient in form, it is still necessary to make the same effort—to spend as much energy in lifting the weight as without the machine. This result was found, on considering various cases in mechanics, to be general: but in the search after an exact conception of this thing that appears to remain unchanged in mechanical actions, a difficulty arose. While it was not possible to get more work out of a machine than was put into it, by any ingenuity of arrangement, actually

machines gave out somewhat less work than was put into them : this loss of work is easily traced to the imperfection of the machines—to friction—being greater, the greater the mechanical imperfection. But if it cannot be traced further the conception of energy is too imperfect to be of much use.

Now it is observed that whenever friction occurs in a machine, heat is produced. This led certain physicists to look for the explanation of the lost work in a production of heat. The experiments of Rumford, about 1800, were strongly in favour of this view ; but it was some forty years later that Joule and others clearly formulated it, and performed exact experiments that led to its universal adoption. It was shown by Joule that when heat is produced by an expenditure of work the quantity of heat produced is in strict proportion to the work spent. Further, that work can be produced by an expenditure of heat is obvious from the action of the steam engine, where the heat of the furnace is the direct cause of the ability of the machine to do work. It needed, then, only the experiments of Hirn, who proved that in a steam engine the work done is in strict proportion to the heat used up, to complete the chain of argument, and show that heat and mechanical work are interchangeable quantities, measurable in a common unit, and may be regarded as merely two aspects of a common quantity—energy. Add to this the fact that many other forms of energy are known, which may be converted into mechanical work and back again—energy of a compressed gas, of a bent spring, of the electric current, of chemical affinity and so on—and we arrive at a conception of energy as a common reality underlying these various phenomena, and binding them together : as a quantity, remaining fixed in amount, merely passing from one of its forms into another, in all the changes of the physical world. This, then, is the law of conservation of energy : to render it precise in detail it will be necessary first to consider the various classes of physical phenomena in detail, so as to know how to measure each kind of energy : but as the law itself is now regarded as axiomatic, and pervades all modern physics, it may well be stated in general terms in the introduction to the subject.

It is possible to proceed in either of two ways : the various kinds of energy and their mutual relations may be classed so far as it is possible to study them by direct experiment, avoiding any



hypothetical suggestions as to their ultimate character, or as to a fundamental identity of structure underlying them ; this is the method of *energetics*. Or the phenomena of mechanics being regarded as the simpler, and the more completely intelligible, the attempt may be made to represent all other phenomena as fundamentally of a mechanical character. This leads necessarily to the hypothetic explanations of the *atomic theory*. Without expressing an opinion as to which method leads furthest down into the reality of natural phenomena, we shall make use of the atomic theory when occasion arises, and in any case begin by dealing with the simple facts of mechanics.

## § 2. Length and Time.

In order to measure the various quantities occurring in dynamics and physics it is necessary to start with one or more units defined in an arbitrary manner : in practice it is found convenient to choose three such—the units of length, time, and mass.

The unit of length now adopted in scientific work is the *centimetre* (cm.): it is defined, in a purely arbitrary manner, as the hundredth part of the length of a certain standard metre rod kept in the archives at Paris. It is true that the metre was intended to be  $\frac{1}{10000000}$  of the quadrant of the earth, and is so approximately : but that is immaterial for the definition ; the length of the metre, and therefore of the centimetre, depends on nothing else than the existence of the standard rod in question, and any actual centimetre scale is a copy, more or less direct, of that rod.

It is, however, desirable to bring this arbitrary length into relation with some well-known natural phenomenon : now the length of a wave of light of some definite kind is a constant and universally accessible quantity, and methods have been worked out for comparing it with the length of a standard scale : choosing the three kinds of light red, green, and blue, given out by incandescent cadmium as the most convenient, Michelson has determined the number of waves of light that would stretch through one metre. The number, which is different for each of the three kinds, is 1,553,164 in the case of the red light (in air at 15° and 760 mm.), and it is believed that this determination is correct to within one wavelength. Thus it is now possible to construct a centimetre scale by optical means, without recourse to any existing scale, with a very high degree of accuracy.

For measuring large and small lengths, the following units are in use :—

Kilometre	= $10^5$ cm. <sup>1</sup>
Metre	= $10^2$ cm.
Millimetre	= $10^{-1}$ cm.
Micron (written $\mu$ )	= $10^{-4}$ cm.
Micromillimetre (written $\mu\mu$ )	= $10^{-7}$ cm.

For a description of the numerous instruments for measuring lengths large and small, for constructing scales, and for comparing scales with one another, reference must be made to the works on practical physics. From the unit of length are derived in a systematic manner the units of area and volume. The unit of area is the square centimetre (sq. cm.); that of volume the cubic centimetre (c. c.). A litre is 1000 cubic centimetres.

The scientific unit of time is the *second*. This, unlike the centimetre, is defined by reference to a natural phenomenon, being  $\frac{1}{86400}$  part of the mean solar day.

It is often necessary, in physical (and physiological) experiments, to measure, and, if possible, to record automatically,

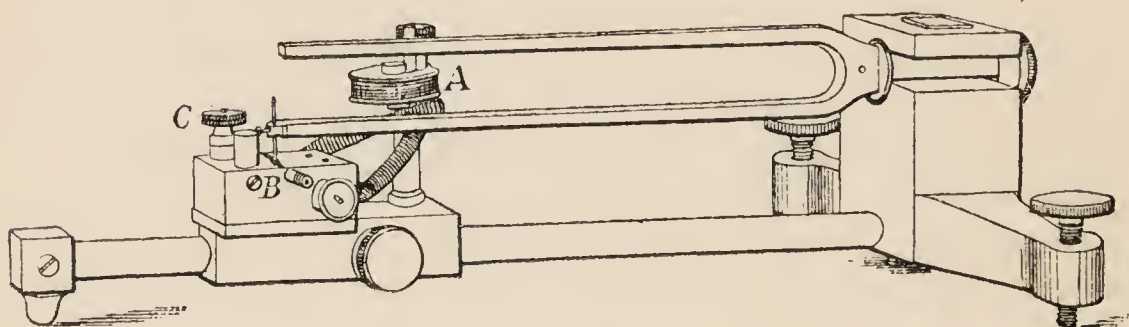


FIG. 1.

periods of time too short for observation in the usual manner by means of a clock. An excellent standard for short intervals of time is a *tuning-fork* whose vibrations are maintained electromagnetically. A short bar of iron, wound with a coil of insulated wire A, constituting an electromagnet (Fig. 1), is fixed between the prongs of the fork : a short piece of platinum wire is attached to one prong in such a manner, that when the prong is pressed

<sup>1</sup> Very large and very small numbers are commonly expressed by means of a power of 10. Thus 5,870,000,000 may be written  $5.87 \times 10^9$ , and 0.0000026 as  $2.6 \times 10^{-6}$ . The rule is, move the decimal point as many places to the right as the index shows if it be positive, to the left if negative.



slightly outwards the free end of the wire dips into a cup of mercury B attached to the framework of the electromagnet, but, electrically insulated from it; one end of the coil is soldered to this framework, the other to an insulated binding screw C. This binding screw and that of the mercury cup are connected to the poles of a battery (say one small accumulator): then if the fork be started vibrating by hand, the prong when it has moved outwards will bring the platinum wire into contact with the mercury, thus completing the electric circuit, and allowing an electric current to flow; this magnetizes the bar of the electromagnet which pulls the prongs together slightly: on their flying inwards again the electric connexion is broken, so that the electromagnet does not resist their further inward movement. A slight impulse is thus given to the fork at each vibration and its vibrations maintained indefinitely. Such forks may make 64 or 100 or even more vibrations per second: they have been employed to drive special clocks indicating hundredths of a second (made by König of Paris), or they may be used in connexion with a chronograph.

A *chronograph* is an automatic arrangement for registering time. Many patterns of such have been designed: Fig. 2 shows one (made by the Cambridge Scientific Instrument Co.) suitable for physiological work. It is driven by the clockwork A which itself derives its energy from a strong spring wound up by the handle B. The clockwork, however, having no pendulum or hair-spring to regulate it, as in a clock or watch, would go faster when the resistance of the mechanism it had to drive was less, and vice versa. To maintain the desired constancy of speed, the fan C is driven at high speed by the clockwork: this suffers so considerable a resistance from the air as to make the resistance of the mechanism small in comparison, and the running is therefore very steady. Moreover, the vanes of the fan are jointed so that the faster they run, the more they tend to fly outwards: hence if the speed gets too high there is more air resistance, and this brings it down to the normal level again: if the speed be too low the vanes drop inwards and offer less resistance, so that the speed rises again. Such a mechanism, which can be designed on several different principles, is called a governor. The clockwork drives the disc D, and against that is pressed the friction wheel E. The latter rotates in consequence of its friction against

the point of the disc which it touches ; but as the points further out on the disc obviously move faster than those near the centre, the friction wheel can be made to rotate faster or slower accord-

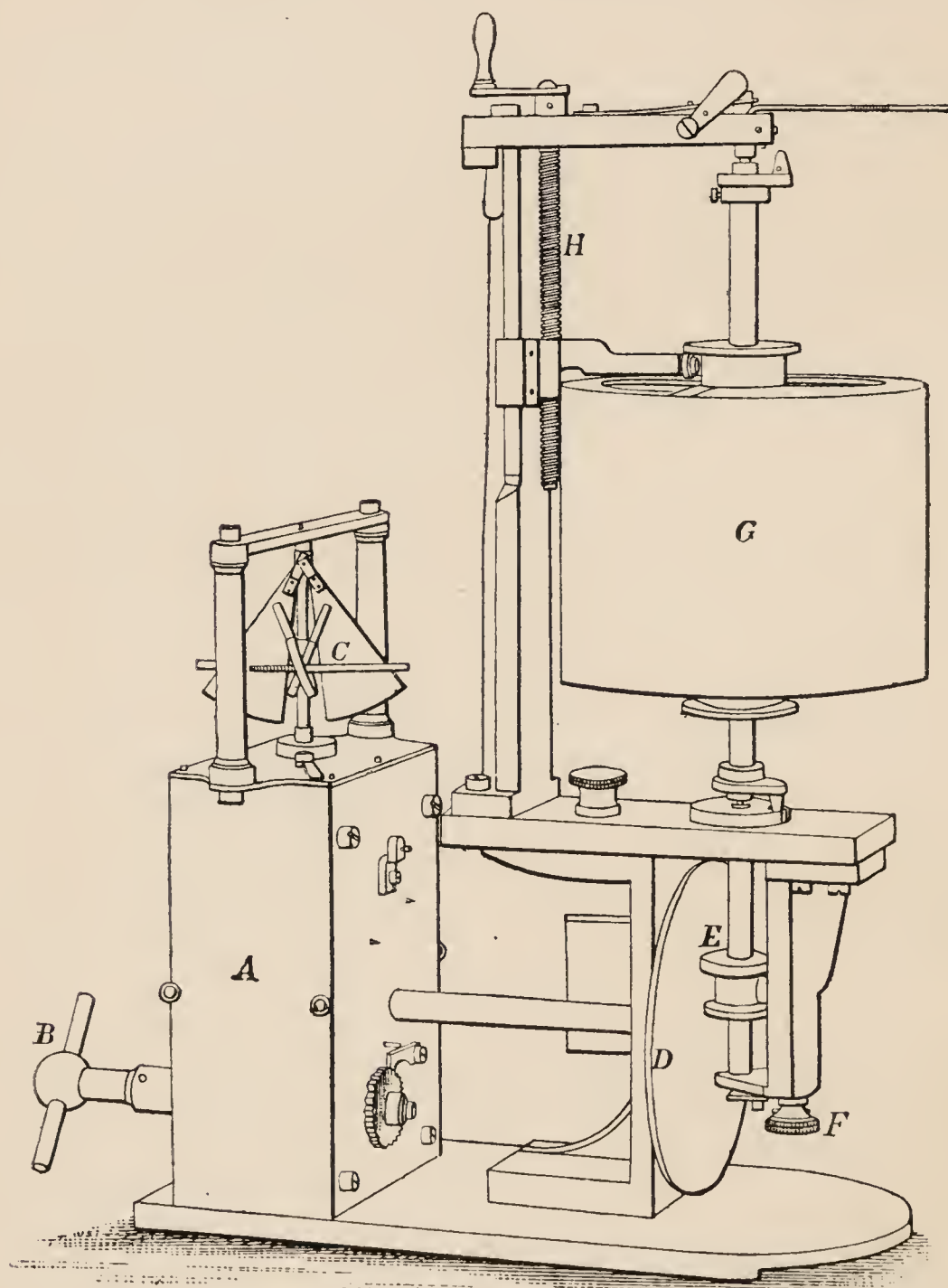


FIG. 2.

ing as it is placed further from or nearer to the centre of the disc : the screw F allows of the vertical motion required to effect this variation in speed. The friction wheel drives the drum G, on which is wrapped a sheet of paper to receive the records. The

drum can be raised or lowered by the screw H without interfering with its rotatory movement, so that the tracing made on the paper, instead of being limited to one rotation of the drum, can be drawn out into a spiral. Many of the details are different in other patterns of chronograph, but the object of all kinds is to secure a uniform rotation of the drum.

Records are made on the moving sheet of paper, by a pen of special construction, or, if the paper be smoked, by a simple sharp point. If the pen be kept still, in contact with the paper, a horizontal line is produced as the paper gradually moves under its point. Any movement of the pen which it is desired to

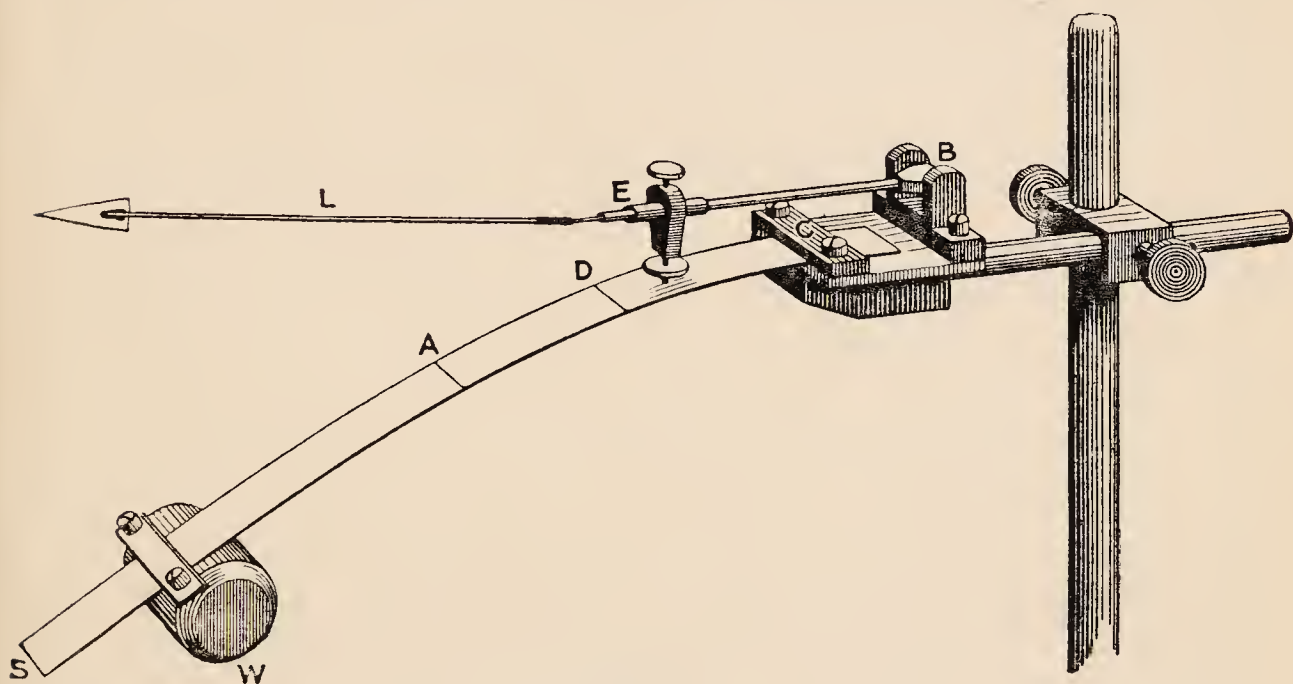


FIG. 3.

record must therefore be crosswise to the movement of the drum. If the pen be sharply pulled up or down and let go again, a notch appears on the curve. This may be accomplished by a time-marker which consists of a light lever carrying the tracing point, arranged to be pulled down by the electromagnet, whenever a current is passed through the latter: the lever recovers by means of the spring. If a tracing point be attached to a regularly oscillating body, such as the tuning fork already described, a wave curve will be drawn by it, which will serve as a scale of times (see e. g. Fig. 9). Another arrangement for the purpose is shown in Fig. 3. Here *s* is a flat steel spring carrying a weight *w* capable of vibrating up and down, and



acting on the tracing lever L. The movements of any other piece of apparatus may be communicated, either by direct mechanical means or electrically, to a similar tracing point, and the indications of the apparatus recorded along with the scale of times.

The notion of *velocity* or speed is derived from those of length and time, and therefore velocities may be best expressed in a unit systematically connected with those of length and time, viz. in centimetres per second: even if a velocity does not remain constant throughout one second, it may still be expressed by measuring the distance traversed in some very short interval of time, and dividing the distance by the time taken. Velocity may usually be regarded as including a conception of the *direction* in which it takes place<sup>1</sup>: and in particular, if movement to and fro along one line be considered, velocity in one sense may be described as positive; in the opposite sense negative.

When the velocity of a body varies, the body is said to have *acceleration*: positive if the velocity be increasing, negative if it be decreasing. Acceleration is measured by the change of velocity occurring in unit time: hence as unit of acceleration we have

$$\frac{\text{unit velocity}}{\text{unit time}} = \frac{\text{cm. per sec.}}{\text{second}} = \text{cm. per sec. per sec.}$$

The most important instance of acceleration is that due to the gravitational attraction of the earth. It is found to be the same on all bodies of whatever size or material, and varies but slightly all over the earth's surface: it amounts to about 981 cm. per sec. per sec., and is directed vertically downwards. Hence a body allowed to fall freely will, at the end of one-tenth of a second, possess a speed of 98.1 cm. per sec., at the end of two-tenths, 196.2 cm. per sec., and so on (ignoring the resistance offered by the air or other medium through which it may be falling).

### § 3. Mass.

The unit of mass adopted in scientific work is the *gram*: this, like that of length, is defined in an arbitrary way as the thousandth part of the mass of a standard piece of platinum

<sup>1</sup> A quantity which implies direction in space as well as magnitude is called a *vector*. Such are velocity, acceleration, force, magnetic moment, and others.



kept at Paris. This standard was intended to be equal to the mass of a cubic decimetre (1000 c.c.) of water at its maximum density. It is not precisely equal, but very nearly so; hence for practical purposes a cubic centimetre of water may be taken as a gram, which is very convenient.

For measuring large and small masses the following units are in use:—

Ton	= $10^6$ grams.
Kilogram (kgm.)	= $10^3$ grams.
Decigram	= $10^{-1}$ grams.
Centigram	= $10^{-2}$ grams.
Milligram (mgm.)	= $10^{-3}$ grams.

The *density* of a substance is the mass of unit volume of it. Hence in the system of units here adopted it is obtained by dividing the mass in grams of a specimen of the substance, by the volume of the specimen in cubic centimetres. From what was remarked above it follows that water has a density of one grm./c.c. The term specific gravity is sometimes used to mean the mass of a body divided by the mass of an equal volume of water (or air): the specific gravity referred to water is consequently almost identical with the density, and the term will not be employed here; but the specific gravity referred to air is sometimes a useful quantity in dealing with gases.

The following table, like others in this book, is intended merely to give the student the necessary general notion of the magnitude of the quantities dealt with. For further information a book of tables must be consulted<sup>1</sup>.

<i>Densities.</i>			
Water	1.00	Crown glass	2.6.
Mercury	13.56	Air	0.0012.
Brass	8.4	Hydrogen	0.00009.

The methods of measuring densities are of considerable practical importance: they all depend on the fact that the mass of a body may be represented by its weight, which is proportional to the mass (see below, § 4).

The most direct way of measuring the density of a liquid is

<sup>1</sup> Such as Everett, *C. G. S. System* (Macmillan, 5s.); Lupton, *Numerical Tables and Constants* (Macmillan, 2s. 6d.); or (much more extensive) Landolt and Börnstein, *Physikalische Tabellen*.

by weighing a small flask full of it, and full of water. The shape shown in Fig. 4 is often used: it is closed by a ground-glass stopper through which a small hole is bored. It is filled with liquid, the stopper then inserted, and the excess which overflows wiped off with a cloth. For accurate work it is necessary to know the temperature at which the flask is full: hence the stopper is sometimes replaced by a thermometer ground to fit the neck of the flask.

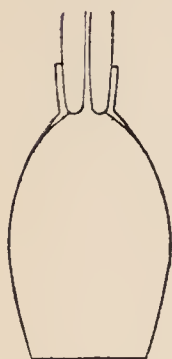


FIG. 4.

More convenient, however, is the Sprengel tube, or pyknometer, shown in Fig. 5; the manipulation of this, which may be taken as typical of that of glass apparatus in general, is as follows: the tube, which may conveniently have a capacity of 5 or 10 c.c., is fitted by *a* to the rubber tube attached to the filter-pump, and, being held slightly aslant, *b* is dipped under nitric acid in a small dish.

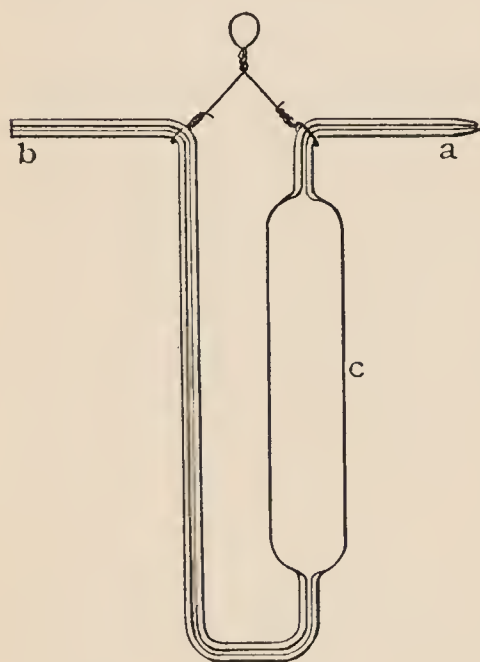


FIG. 5.

On turning on the pump the acid is sucked in; after cleansing, it may be poured out through *b*, if the tube have a bore of about 1 mm. The cleaning may be repeated with various liquids, e.g. nitric (or chromic) acid, water, caustic soda, water, distilled water: then, still by means of the filter-pump, air may be drawn through, whilst the tube is gently warmed over a gas-burner, and in a few minutes it will be dry: *b* should be covered with cotton-wool during this operation to keep dust out. In order to get accurate weighings, the outside should then be wetted and wiped dry with a hard cloth (handkerchief) that does not leave fluff behind; if this is not done the surface retains a variable amount of moisture.

The pyknometer may now be weighed empty: it is then filled with distilled water in the way already mentioned, and when full slipped off the rubber tube and placed upright in a bath of water alongside a thermometer; in two or three minutes it acquires the temperature of the bath, and the quantity of liquid in it is then adjusted to a mark etched on the tube *b*. The capillary effect of the narrow end of *a* keeps *a* full, so that if a piece of filter-paper be held against *a* the water will be withdrawn from *b*. If there is too little water, a drop on the end of a glass rod is held against the end of *a* and water will flow into *b*. The pyknometer is then taken out, wiped dry, and

weighed: the same process is repeated with the liquid to be measured. Evaporation does not take place rapidly from the small exposed surface of a Sprengel pyknometer, so that it may be used for water and aqueous solutions without any trouble. For more volatile liquids glass caps ground to fit the ends serve to prevent evaporation.

If  $W_0$  be the weight full of water less the weight of the empty tube, and  $W$  the weight full of liquid less the weight of the empty tube, the density is  $W \div W_0$ . If an accuracy of 1 in 1000 or better is aimed at, it is necessary to make a correction for temperature: the result, as found, really gives the specific gravity of the liquid by reference to water at the temperature of the experiment: now the density of water (gms. per c.c.) varies a little with temperature: if, then, at the temperature of the experiment it be  $D_0$  (found from published tables), the true density of the liquid, at the same temperature, is

$$= \frac{W D_0}{W_0}$$

When a lower degree of accuracy suffices (say 1 in 200) and a good supply of liquid is available, the common hydrometer is more convenient. This instrument consists of a glass bulb (Fig. 6), weighted so as to float upright, and carrying a graduated vertical stem. When placed in a liquid it sinks till the weight of the liquid it displaces just balances its own weight (see § 8). It does not need to sink so far in a dense liquid as in a light one to satisfy this condition, and hence the point to which it sinks measures the density of the liquid, according to a scale previously marked on the instrument. Hydrometers with scales of various ranges may be bought, and their accuracy checked by one or two measurements with the pyknometer.

Sometimes it is convenient to test the density of a liquid by means of a set of glass beads of varying density. These, which have their densities marked on them, are dropped into the liquid till one that just floats and one that just sinks are found: the density of the liquid then lies between the two.

When only a small quantity of the liquid is available—e. g.

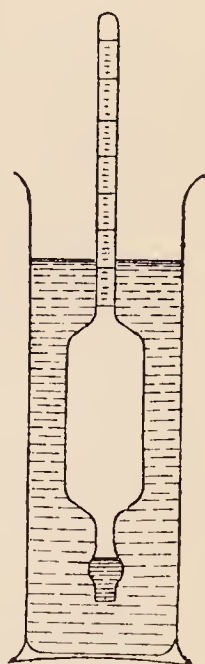


FIG. 6.



blood—the best plan is to prepare a series of liquids of known densities and see in which of these it floats or sinks. The standard liquids may be mixtures of glycerine and water, contained in test tubes. The blood is contained in a capillary pipette, and a small drop of it is pressed out into the middle of such a tube; if it rises, it is lighter than the mixture, and the next lighter mixture must be tried, till it has been found between which pair its density lies.

The density of a solid which occurs in good sized pieces may be found by the method of the hydrostatic balance (areometric method): the specimen is weighed in the usual way ( $W$ ): it is then hung by a very fine silk thread or platinum wire from an arm of the balance in such a way that the specimen is completely immersed in a beaker of water, the weight of the water being, however, borne by a support independent of the balance: it is then found that the weight appears less ( $W'$ ). The density (see § 8) is

$$\frac{W}{W - W'} \times D,$$

where  $D$  is the density of the water or other liquid in which it is hung. If water attacks or dissolves the substance in question, another liquid of known density may be used instead. The method is not applicable to solids lighter than the liquid used, except with the complication of a sinker.

For light solids, however (i. e. most solids other than the metals), especially when occurring in small pieces, the most accurate and convenient method is to find a liquid mixture in which the pieces neither float nor sink, as described above for blood.

The density of gases is found by methods essentially similar to those for liquids: for details, larger works should be consulted.

#### § 4. Momentum: force.

The next dynamical quantity to claim attention is *momentum*, expressively called by Newton the 'quantity of motion.' A moving body is said to possess a momentum which is measured by the product of its mass and its velocity: thus the effective quantity of motion of a body may be large either on account of its having a large mass (e. g. a wagon) or large velocity (bullet). Momentum,

like velocity, is to be considered as having a direction, and consequently being positive when the body is moving in one sense, negative in the other.

Newton's *third law of motion*—the equality of action and reaction—may best be expressed in modern language in terms of momentum. It is that *in any action between two bodies, the momentum gained by the one is equal to that lost by the other*. In interpreting this, it must be borne in mind that momentum has a positive or negative sign, otherwise the statement will be reduced to nonsense. It may be illustrated by a numerical example; thus, suppose a tennis racquet of mass 290 gms. moving with a speed of 400 cms. per sec. to collide with a ball of 60 gms. moving in the opposite direction at 700 cms. per sec. We will take the direction of motion of the racquet as positive: then the racquet possesses, before the collision, a momentum  $= 290 \text{ gms.} \times 400 \frac{\text{cm.}}{\text{sec.}} = 116,000$ , the ball  $60 \times (-700) = -42,000$ .

This information is not in itself enough to determine the motion: it would be necessary to know the elastic properties of the racquet and ball as well: if however we find by observation what happens to one of them, the law of motion will show what must happen to the other. Thus, suppose the ball is driven forward at 600 cms. per sec.; it now possesses a momentum of  $60 \times (+600) = +36,000$ . The momentum has been increased by the blow from  $-42,000$  to  $+36,000$ , a total *increase* of 78,000; hence the momentum of the racquet must have *decreased* by the same amount, and become  $116,000 - 78,000 = 38,000$ ; the racquet is therefore still moving forward (its momentum is positive), but its speed has fallen to  $38,000 \div 290 = 131$  cms. per sec.

The law of equality of action and reaction really serves as a means of measuring masses. We have, it is true, defined a kilogram as being equal to the platinum standard at Paris: but we have provided no means of testing the equality. We now see that if two equal masses collide or otherwise interact, the gain of velocity by the one must be equal to the loss by the other, since in that way the gain of momentum by the one will be equal to the loss of momentum by the other. A means of testing this experimentally is given by Hicks's 'ballistic balance,' which is a device for projecting two bodies against each other with measured velocities and for measuring the velocities with



which they rebound<sup>1</sup>. Thus suppose we have a standard kilo and another body supposed to be equal to it in mass (it may differ in material, shape, and size): project the two in opposite directions with equal velocities, and (for convenience) arrange by means of automatic clips that they do not rebound after meeting. Then if they are really equal they will stop dead after the collision: for beforehand the one possessed a certain positive momentum, the other an equal negative amount. The first body loses all its positive momentum, and the second therefore necessarily gains the same amount, which just serves to neutralize the negative momentum it had, and bring it to rest. If however the masses are not equal, the greater mass will have the more momentum (since the velocities are equal) and the two bodies after collision will move on together slowly in the direction in which the greater mass was moving.

Standards of mass are not actually tested in this way, but by means of their weight: nevertheless the method is important because it shows that the conception of mass is intelligible and complete without introducing that of weight.

The conception of force may be approached either from that of momentum or of acceleration. According to the former method, which agrees with the order of reasoning expressed by Newton's laws of motion, force may be defined as the *rate of change of momentum*: a force can be measured, therefore, by observing the amount of momentum it generates in a measured time, and dividing by that time. Hence as unit of force we have

$$\text{force} = \frac{\text{momentum}}{\text{time}} = \frac{\text{gm. cm. per sec.}}{\text{sec.}} = \text{gm. cm. per sec. per sec.}$$

The expression 'per second' has to be repeated here, as in the precisely analogous case of acceleration. As this unit is important, and its systematic name is lengthy, a special name has been given to it, viz. the *dyne*.

The conception of force, which may be regarded as a matter of common knowledge, derived in the first place from the sense of muscular effort, is rendered scientifically accurate by considering the first and second laws of motion. The first law is that '*Every body continues in its state of rest or of uniform motion in a straight line except in so far as it is made to change that state*

<sup>1</sup> For details see W. M. Hicks's *Elementary Dynamics*, p. 24.



by external forces.' So long as the body is at rest, or moving with uniform velocity, it possesses a constant momentum ; hence the *effective force* on it (measured by the rate of change of momentum) is zero. To say that the effective force on a body is zero, means one of two things, either (1) that there are no forces at all acting on it (this never actually occurs in nature), or (2) that two or more forces act on it, but in such a way as to neutralize each other's effect, i. e. produce *equilibrium*. A body may therefore be in equilibrium, and yet in motion, provided the motion be uniform and in a straight line. Thus if a heavy body be placed on the floor, it is subject to two forces, viz. the attraction of the earth, downwards (i. e. its weight), and an equal pressure from the floor, upwards : these two neutralize, and the body remains at rest. If, however, the body be placed on the floor of a lift, which is ascending with uniform speed, there is again no change occurring in its momentum, therefore no effective force on it, so that again the pressure of the floor must be just sufficient to balance its weight. But if the lift be starting to go up, and consequently suffering an acceleration, the case is different : the body is then gaining momentum, so that an effective force is acting on it, which shows that the pressure of the lift floor must be different from (in fact greater than) the weight, and the two forces are no longer in equilibrium. When the several forces acting on a body are not in equilibrium the excess which remains in some direction is called the *resultant* of the forces, and constitutes the effective force producing change of momentum.

It is to be noted that since the notion of momentum (or of velocity) includes that of direction in space, a change of direction is a change of momentum, and therefore requires a force : thus it is only uniform motion in a straight line that implies absence of force : uniform motion, e. g. in a circle (stone tied to a string and swung round uniformly), requires a force constantly acting (tension of the string).

The second law of motion may be stated as follows : *Rate of change of momentum is equal to the force applied, and takes place in the direction in which the force is applied.* This, it will be seen, includes the definition of force already given : it also gives information as to the resultant of several forces, as will be shown in detail below.

A little consideration of the above statement, or of the nature

of the unit of force, will show what is the relation between force and acceleration: the unit of force, the dyne, or gm. cm. per sec. per sec. may not only be regarded as (gm. cm. per sec.) per sec., i. e. as momentum divided by time, but as gm.  $\times$  (cm. per sec. per sec.), i. e. as mass multiplied by acceleration. A force therefore may be considered as measured by the product of the mass it is acting on, into the acceleration which it produces on that mass. The cross relations between these quantities may in fact be simply illustrated by a diagram:

velocity	momentum
acceleration	force

Thus each of the two lower quantities is the time-rate of change of the quantity standing above it; while each of the quantities on the right is identical with the corresponding quantity on the left, multiplied by mass.

One of the most important deductions to be drawn from the laws of motion is as to the property known as *inertia*. When a mass at rest is set in motion, an acceleration is produced in it, and therefore a force has to be exerted, independently of any frictional or other resistance to the motion. This inertia, opposing the starting of a heavy body (opposing equally its stoppage), is a familiar fact, evidenced by the behaviour of carts, bicycles, &c.: and it is nothing more than the essential meaning of mass as shown by the second law of motion. A particular

case, important to the physiologist, will serve to render the notion of inertia more definite.

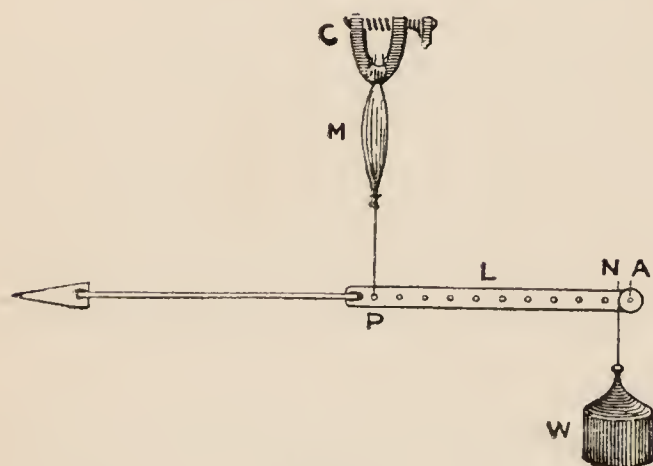


Fig. 7.

To study the contraction of muscles, an arrangement like that of Fig. 7 is used: the muscle, fixed at the upper end, is attached below to a lever, hinged at A, so that when the muscle is stimulated the tracing point at the other end of the lever is lifted

and makes a corresponding mark on the chronograph paper which it touches. If a mass of say  $m$  grams is hung at P immediately below the muscle, when the contraction occurs, an upward acceleration has to



be communicated to it: the tension of the muscle, which when at rest was equal to the weight of  $m$  grams, is consequently now somewhat greater. On the other hand, when the lever has risen nearly as high as it will go, and is coming to a stop, its inertia tends to carry it on: in other words, it is now suffering a negative acceleration (its velocity upwards is being lessened) and the tension of the muscle is somewhat less than the weight. This fluctuation in tension causes an oscillation to be set up, and recorded on the chronograph, as may be seen on comparing Figs. 8 and 9 (taken from Brodie's *Experimental Physiology*). Each represents a simple muscle twitch, Fig. 8 with a so-called 'light' lever possessing little inertia: Fig. 9 with a heavy lever; the latter clearly shows an oscillation of the writing point which is not proper to the muscular movement studied, but introduced by the apparatus. To avoid this misleading effect, it is not however desirable to use a very light lever unloaded, since this would not keep the muscle stretched: the way out of the difficulty was found by Fick, and is known as the isotonic method. It is illustrated in Fig. 7, and consists in hanging the load much nearer to the axis, and increasing it in proportion: thus instead of  $m$  grams at  $P$  we may, as in the figure, use  $20m$  grams at a point  $N$   $\frac{1}{20}$  of the distance from the pivot  $A$ : this produces the same tension in the muscle (see § 6 on moments); now the movements of  $N$  will only be  $\frac{1}{20}$  as small as the corresponding movements of  $P$ , and the acceleration therefore reduced to  $\frac{1}{20}$  that of  $P$ . The mass moved is 20 times as great, and since the force exerted is measured by the product of mass and acceleration it is the same in actual amount as when the small mass is hung at  $P$ . What we are concerned with however is the relative fluctuation in the tension: and as the weight at  $N$  is 20 times as great as that required at  $P$ , the same actual force introduced by inertia will be of far less consequence relatively to it.

Of all forces, the most frequent and important in practice is that of gravity, or *weight*. We have seen that the attraction of the earth produces on all bodies near its surface an acceleration of 981 cm. per sec. per sec. The acceleration varies slightly from one place to another on the earth, but at any one place it is precisely the same on all bodies. Hence at any one spot, a gram of matter has a weight of (about) 981 dynes, and each gram has precisely the same weight as each other gram, regardless of its quality, size, or shape. It follows that two masses can be compared by comparing their weights, i. e. by weighing one against the other in an ordinary balance, as is commonly done, without having to fall back on the much less accurate and convenient methods depending on inertia, such as that of

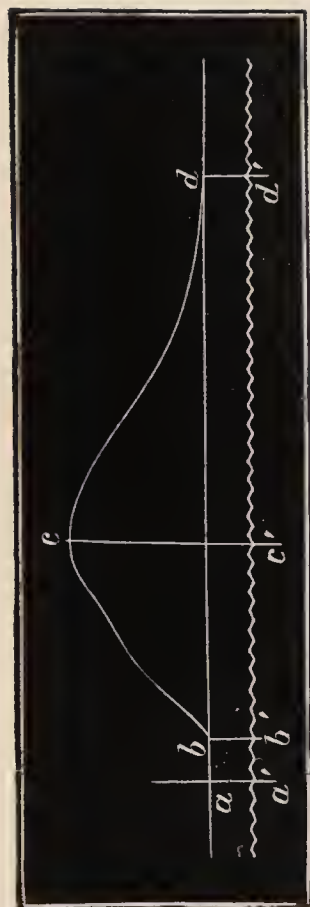


FIG. 8.

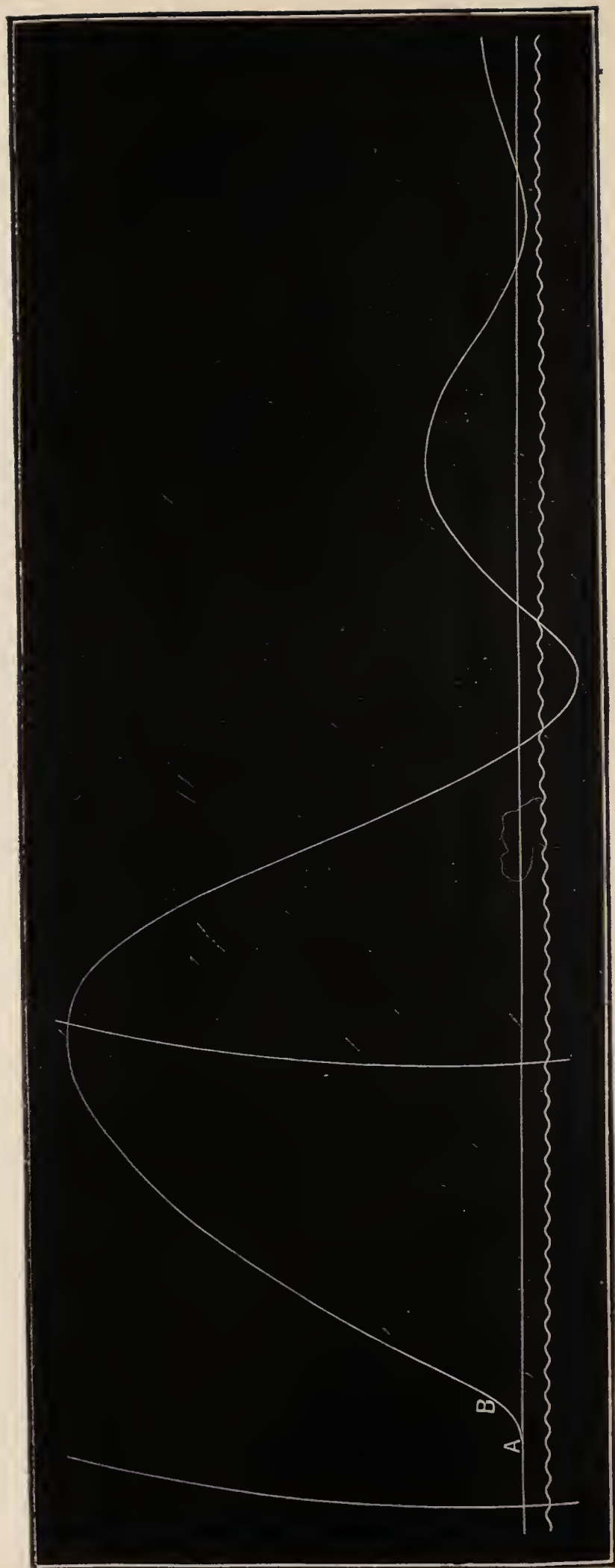


FIG. 9. Time scale : 1 complete vibration =  $\frac{1}{100}$  sec.

the ballistic balance. But it must be borne in mind that the fact that the weights of bodies are proportional to their masses is only an observed result, although observed for all known kinds of matter, and is not an essential part of the meaning of mass. So far as we can at present see it would be quite possible to conceive of some kind of matter obeying the second law of motion (possessing 'mass' or 'inertia') and yet not suffering gravitational attraction (i.e. possessing no 'weight'), just as a piece of copper and a piece of iron may have the same mass, yet one is strongly attracted by a magnet, the other not at all.

Owing to the frequency with which weights have to be dealt with in practice, it is often convenient to express all forces by means of the weights they are equal to: thus we may not only speak of the 'weight of a gram,' meaning the attraction of the earth for a gram mass, and therefore 981 dynes, but we may describe some other force, such as the tension of a horizontal string, which has nothing to do with gravity at all, as being a 'force equal to the weight of  $x$  grams,' or for brevity 'a force of  $x$  grams.' A force so expressed may always be converted into dynes by multiplying by 981.

In a curve such as Fig. 9 distances from the horizontal line express displacement of the end of the muscle from its normal position. Hence to measure the rate at which the muscle is moving at any instant we must take the height above the horizontal line at the beginning and end of some short interval—say one vibration of the tuning-fork. When the muscle is moving at a constant rate, the line traced out rises equal amounts in equal intervals of time, and is consequently straight (but not horizontal); and the steeper it is the greater the distance travelled in a given time, i.e. the greater the velocity. If Fig. 9 be examined it will be seen that the muscle is at rest up to a certain point A; that a little later, at B, it is moving uniformly (the tracing being straight), and from measurement of the curves the velocity then will be found about 20 cm. per sec. During this period then the muscle is exerting such a force that its movable end, to which the weight is attached, is accelerated upwards, increasing in speed from 0 to 20 cm. per sec. in the course of 0.01 sec.; accordingly the measure of the acceleration is  $20 \div 0.01 = 2000$  cm. per sec. per sec. This involves a force which is in addition to the force required to balance the weight: the total tension is therefore 2981 dynes on each gram suspended, and if we say there are  $x$  gms., we may write the equation (expressing the second law of motion) thus:—



Tension of muscle ( $2981 \times$  dynes upwards) — weight supported ( $981 \times$  dynes downwards) = effective force ( $2000 \times$  dynes) applied to the lever, and consequently causing an acceleration ( $2000$  cm. per sec. per sec. upwards) in the latter.

### § 5. Work : energy : power.

Mechanical work is done when a body is moved in opposition to a force that resists the motion. The measure of the work done is consequently the product of two factors: (i) the force overcome, (ii) the distance through which the body is moved against it. The simplest case we can consider is that in which the movement takes place in the line of action of the force: e. g. the line of action of weight is vertically downwards, so if a weight be lifted vertically upwards, work is done on it to an amount equal to the product of the weight (force) into the height. The natural measure of work on the C. G. S. system of units, therefore, is the dyne multiplied by the centimetre, and to this unit, on account of its importance, a special name has been given—the *erg*. If now the movement do not take place along the line of action of the force, the work is measured by multiplying the force by the *distance moved in the direction of the force*, e. g. if a weight be lifted slantways, the vertical distance only must be considered in estimating the work: no more work is done in lifting a weight on to a table 80 cm. high slantways, although the space traversed may be a metre or more, than if the weight be lifted vertically up to the table. It may even happen that a force is acting on a body and yet not moving it in its own direction at all: this is the case with a body moving uniformly round a circle (p. 33). Thus if the moon travels in a circle round the earth, it is because of the pull exerted on it, gravitationally, by the earth, yet as the pull is always in the direction joining the earth and the moon, and the moon does not move along that line, either towards the earth or away from it, but at right angles across the line, no work is done.

*Energy* is the capacity for doing work, and is therefore of course to be measured in the same unit, the erg. As stated in § 1, the fact that numerous forms of energy are known, and that they can be converted into one another without change in amount, is the leading principle of modern physics. These

various forms will be dealt with in turn later : here we need only consider a certain preliminary distinction into two kinds, and the relation between the two as it occurs in abstract dynamics.

A body in motion has a capacity for doing work, in consequence of its motion : this is known as its *kinetic energy*.

Also a body (or system of bodies) usually has a capacity for doing work due to its *configuration*, i.e. the relative position of its parts : this is known as its *potential energy*, e.g. a spring when compressed has potential energy, since on changing its configuration, viz. on reverting to its normal length, it can do work.

When force is exerted on a body, unless it be balanced by other forces, it is spent, we have seen, in giving acceleration to the body : the same fact may be stated in terms of energy, thus : when work is done on a body, unless it be balanced by work done by the body on others, it is spent in imparting velocity, and therefore kinetic energy to the body. Since by the law of conservation the kinetic energy generated must be equal to the work spent, we may learn from this how to measure kinetic energy. For this purpose we may take the most familiar case, that of a falling weight : here the 'system' consists of the weight and the earth : the change in configuration is the change in distance between them : the system possesses potential energy of gravitation, which is spent as the weight falls towards the earth. If now the fall of the weight is caused, by means of a string, to drive a piece of machinery, such as a clock, work is done by the weight on the clock, so that very little of the work spent by the weight goes to produce kinetic energy in itself, and its rate of descent is very slow. If however the weight fall freely, so that there is no force resisting it, and no work done by it on external bodies, then all the potential energy of gravitation that it loses reappears in the form of kinetic energy.

Suppose, for simplicity, that the falling mass be one gram, and that it falls for  $t$  seconds : let  $g$  be the acceleration of gravity (about 981) ; then it follows from the definition of acceleration that the velocity acquired at the end of the time is

$$v = gt.$$

But since the velocity is increasing uniformly (the acceleration



is constant) its average value during those  $t$  seconds is the mean between 0 and  $v$ , i. e.  $\frac{1}{2}v$ , so that the distance through which the weight has fallen,  $h$ , is

$$h = \frac{1}{2}vt.$$

On the other hand, the force with which the earth is attracting the mass downwards (the 'weight') is  $g$  dynes: and therefore from the definition of work, the work done by the earth on the weight (or potential energy lost by the system: earth + weight) is  $gh$  ergs; this then must be the kinetic energy acquired by the falling mass. But

$$gh = g \times \frac{1}{2}vt = \frac{1}{2}v^2,$$

i. e. the kinetic energy of the moving gram is measured by half the square of its velocity: the kinetic energy of any moving body is measured by *half the product of its mass into the square of its velocity* ( $\frac{1}{2}mv^2$ ).

We saw, above, that it is common, and in many cases convenient, to measure forces in terms of weights, instead of in the systematic unit—the dyne: when that is done, it is also convenient to measure work in a corresponding gravity unit: this is the *gram-centimetre*, i. e. the work done in lifting a gram weight through a height of one centimetre: since the weight of a gram is 981 dynes, the work done in lifting it one centimetre is 981 ergs.

It must be carefully borne in mind in the calculation of kinetic energy given above, that the kinetic energy of a body has nothing to do with the attraction of the earth: that its expression ( $\frac{1}{2}mv^2$ ) gives the amount of it naturally in the systematic unit of energy, and that if some other quantity of energy be given in gravity units it must first be reduced to ergs before a comparison can be made. If the systematic units of force and energy be used throughout, no complication will arise.

Even the gram-centimetre is an inconveniently small unit in practice, especially for engineering purposes. The following table will sufficiently explain the other units in use:—

Gram-centimetre =	981 ergs.
Kilogram-metre =	98,100,000 ergs.
Joule =	10,000,000 ergs.
Therm (calorie) =	42,000,000 ergs.

The first two are gravity units, the second being that commonly

adopted by continental engineers: the third is designed as a convenient multiple of the erg, and is much used for scientific and especially electrical purposes: the fourth is the unit of heat, and since, as stated in § 1, heat is a kind of energy, it is also a unit of energy.

In dealing with momentum it was pointed out that that quantity must be understood as being associated with a particular direction in space, and that if a body travelling in one sense along a line is considered to have positive momentum, then in travelling along the same line in the opposite sense its momentum is negative. With this understanding, the third law of motion may be regarded as a law of conservation of momentum, since, one body losing as much momentum as the other gains, the total amount remains unchanged. But the law of the conservation of energy is not to be understood in this sense: *energy* is not a vector: it *has no sign*: a body cannot have a negative amount of energy, either potential or kinetic: such a statement would have no meaning. Hence the law of conservation implies that the actual arithmetic total of energy possessed by any system of bodies remains unchanged, whatever action may take place amongst those bodies, provided no energy is put into the system from outside.

When a force doing work does not remain constant in amount, it becomes necessary to consider what is the right way of averaging the force in order to calculate the work done. This is best done graphically, i. e. by a method of representing the quantities considered, by lines on a diagram, somewhat in the fashion already made use of in connexion with the chronograph. In Fig. 10 let the vertical distances (axis of  $y$ ) represent the magnitude of a force: whilst horizontal distances (axis of  $x$ ) represent the

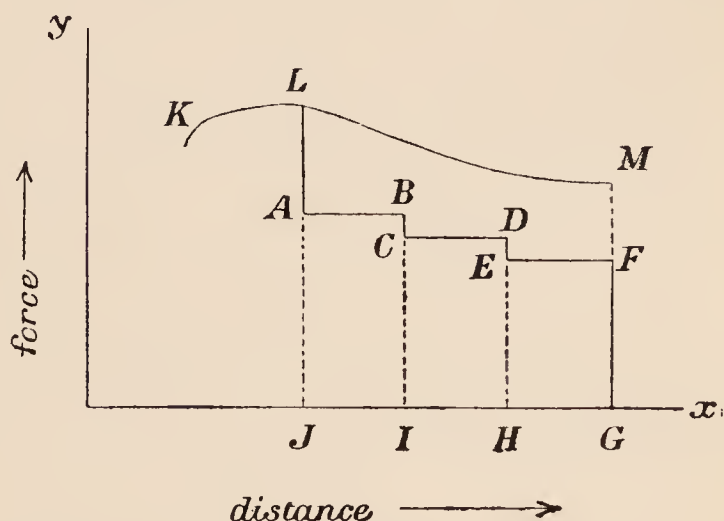


FIG. 10.

distances traversed by the point of application of the force (in the direction in which the force acts). As the most familiar of such



diagrams are the 'indicator diagrams' used to show the action of a steam engine, it will help to a precise understanding, if we suppose Fig. 9 to be such. The problem, there, is to find the amount of work done by a piston in its 'stroke,' i. e. its movement along the cylinder of the engine; hence the horizontal distances in the diagram must represent the length of the cylinder along which the piston travels, while the vertical distances represent the force with which the steam is driving the piston at any moment of its stroke. Now suppose that whilst the piston is moving from  $J$  to  $I$  the force on the piston remains constant and equal to  $AJ$  or  $BI$ ; then from the definition of work (force  $\times$  distance) it follows that the area  $ABIIA$  is in proportion to the work done by the steam during this movement. Similarly, if whilst the piston moves from  $I$  to  $H$  the force of the steam remains constantly equal to  $CI$  or  $DH$ , an amount of work measured by the area  $CDHIC$  is done: and this area, it will be seen, adjoins the preceding area. Now the force does not actually change in this discontinuous manner, but by making the sections  $JI$ ,  $IH$  &c. during which the force is regarded as constant small enough, we may approach as nearly as we please to the actual continuous change such as is represented, e. g. by the curve  $KLM$ . In this way we see that if the force at any moment be represented by the height of the curve  $LM$ , the work done by it is represented by the area enclosed by the curve, two perpendiculars and the axis of  $x$ ; e. g. the work done whilst the piston moves from  $J$  to  $G$  is represented by the area  $LMGJL$ . In engineering practice the curve is drawn automatically by means of a piece of apparatus called an indicator: in this, as in the chronograph, a pencil is moved vertically whilst a sheet of paper wrapped on a drum is moved across it; the vertical movement is given by the steam pressure, acting against a spring, and consequently measures the force exerted by the steam; the horizontal movement is not, as in the chronograph, caused by clockwork, since the intention is, not to record equal times, but equal spaces transversed by the piston: the drum is, therefore, actuated by a cord directly connected to the piston. When the diagram has been drawn, the sheet of paper may be detached and the area measured, in order to calculate the work done in the stroke.

The word *power* (or 'activity') as used in dynamics means rate of doing work; it is therefore to be measured by the number of ergs done per second. Two other units are in common use:—

$$\text{Watt} = 10,000,000 \text{ ergs per sec.} = 1 \text{ joule per sec.}$$

$$\text{Horse-power} = 7,460,000,000 \text{ ergs per sec.} = 746 \text{ watts.}$$

As an instance of calculations of work and power, we will take the action of the heart. This is estimated to discharge



180 gms. of blood at each beat, against an excess of pressure in the aorta equal to one-third of the atmospheric pressure. This (p. 37) is equivalent to lifting the blood up to such a height as would suffice, in a column of blood, to produce a pressure of one-third atmo, or about 325 cms. Hence the work done at each beat is  $180 \times 325 = 58,500$  gm. cm. If there be 72 beats per minute this is  $58,500 \times 72 = 4,212,000$  gm. cms. per min. or 42.12 kilog. metres per min. or  $42.12 \times 9.81 = 413$  joules per minute. The action of the heart is of course not uniform in a second, but this is equivalent to an *average activity* of 6.9 joules per sec. or 6.9 watts. This is for the left ventricle only. For comparison it may be mentioned that 40 or 50 watts is as much as a man can do in the way of muscular work, continuously.

The relations of energy and power to the other dynamical quantities may be conveniently shown by a diagram thus:—

space rate $\rightarrow$	
time rate $\downarrow$	—
	momentum
	force
	—
	power (activity)

Here each quantity is the rate of change with the time, of the quantity immediately above it: power being rate of change of energy; force, rate of change of momentum. Again, the right-hand column is derived from the left, by being its rate of change in space, or the left can be derived from the right by multiplying it by a length: this is shown in the relations between energy and force; energy (work) is obtained by multiplying force overcome by the length through which the point of application of the force is moved: conversely, force is the rate at which the energy is changed (or work is done) when the point of application is moved by unit amount<sup>1</sup>.

<sup>1</sup> The diagram might be extended by putting in the left-hand top corner the quantity known as 'action'—this quantity, however, is never used in elementary dynamics.

## § 6. Vectors.

We have already had occasion to remark on the difference in character between energy as a quantity, and certain other quantities such as momentum: they are not to be treated by the same mathematical rules, on account of the association of momentum with a direction in space, which is foreign to the notion of energy. Dealing more explicitly with this distinction we find that there are three kinds, or stages, of quantity, which may be described as (i) *arithmetic*, (ii) *algebraic*, (iii) *geometric*. The first, which includes energy, is the simplest kind of quantity, involving no conceptions of either sign or direction: it is that treated by the ordinary rules of arithmetic, so that to add two quantities of energy together is an operation in the sense of simple addition: if a body contains three ergs of energy, and two ergs are imparted to it, then it possesses five—the sum is necessarily greater than either component. To gain the conception of algebraic quantity it is necessary to supplement this by positive and negative sign: examples of such quantities, involving sign but not direction, are to be found in Physics, but for the sake of familiarity we will choose, rather, one from common life, viz. money. If money owned be described with the positive sign, then money owed is negative; and it is not only possible to add together two positive amounts, but also to add up various sums owned and owed, giving a negative sign to the latter, and so arrive at the total value of one's property. In this case of course the sum is not necessarily greater (in the arithmetic sense) than the parts: thus if one has £3 in one's purse, £40 at the bank, and owes one's tailor £15, to arrive at the total value it is necessary to perform the mathematical operation represented by

$$\begin{array}{r}
 + 3 \\
 + 40 \\
 - 15 \\
 \hline
 + 28
 \end{array}$$

The example given on p. 13 with regard to momentum is of that kind: length, velocity, acceleration, momentum, and force, are all quantities which are to be reckoned positive when in

one sense, negative in the opposite sense, and so possess the characteristics of algebraic quantity. It is true they go beyond this, and being associated with direction in space, are really geometric; but if for the moment we neglect this fact and consider a length, momentum, &c. in one line only, we get a good example of the kind now considered. Thus in Fig.

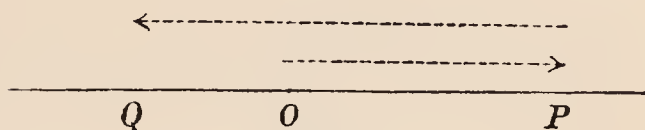


FIG. 11.

11, taking  $O$  as the starting-point, the distance  $OP$  may be regarded as a *step* in one sense along the line (to the right). If we regard this as positive, then any step made in the other sense (towards the left) must be regarded as negative: such e.g. is  $PQ$ . Hence the result of starting from  $O$  and adding together the two steps  $OP$  ( $= +2$  cm.) and  $PQ$  ( $= -3$  cm.) is to arrive at  $Q$ , and the sum of the two steps is  $OQ$  ( $= -1$  cm.).

These results may seem almost too obvious to need explicit remark; but that is certainly not the case as regards the addition of geometric quantities, or vectors. Here again a length, as the simplest of vectors, may best be taken as example, and a length may be looked upon as a step. In Fig. 12  $OP$  is not only a length of two centimetres; it is that length in a direction (which we will call *ENE*). To this is to be added the length  $PQ$ , which is three centimetres, but in the north-west direction. The result is to arrive at the point  $Q$ ; giving the step  $OQ$  as the sum of the other two.

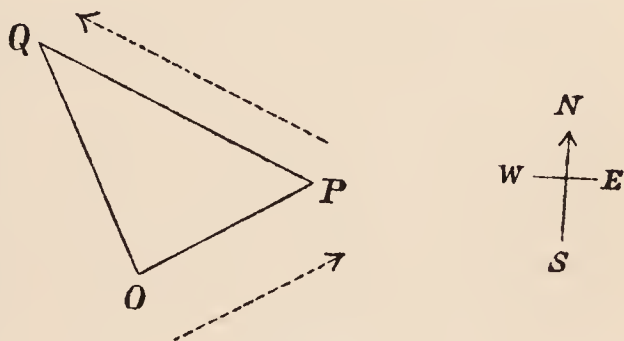


FIG. 12.

This is the proposition known as that of the triangle of forces, for all vectors, including of course forces, can be added in this way. If it be desired to find the resultant (i.e. the sum) of two forces acting at a point, this can be done by drawing a diagram in which the forces are represented by lines drawn parallel to the forces and of lengths to represent their magnitude on some fixed scale.

These principles, especially as regards forces, can be illustrated from mechanics, and as well from the mechanics of the human body as anything. Thus the radius and the flexor muscles of the fore arm constitute a sort of bracket for supporting the hand. The parts they play are similar to those of the strut  $A$  and tie  $B$  in Fig. 13. It is easily seen that the strut is in compression, and consequently exerts a thrust  $x$  at  $L$  (if it



were not strong enough it would crush up), whilst the tie is in tension and exerts a pull  $y$  at  $L$  (if it were not strong enough it would be pulled

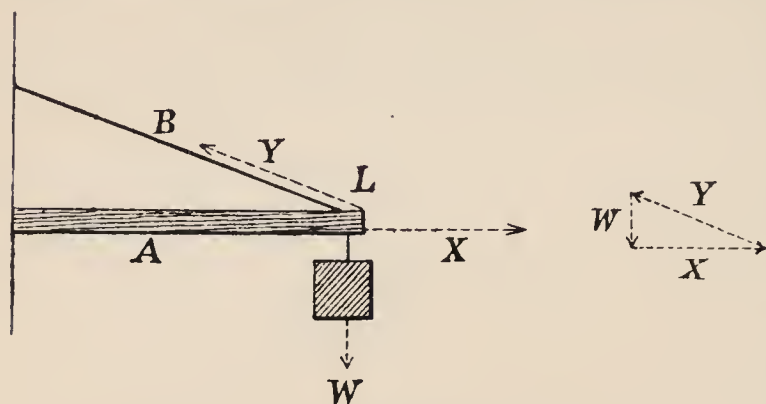


FIG. 13.

apart). These two forces when added together (as vectors) have a resultant which just balances the weight  $w$ . This is shown graphically in the triangle at the side. The resultant of  $x$  and  $y$  would be the third side of the triangle, directed upwards, and it produces equilibrium with the weight which is repre-

sented by that line drawn downwards. It should be noticed that in this example both forces  $x$  and  $y$  are considerably greater than their resultant—the reason being that they so nearly oppose each other. In the figure as drawn, if the weight were 5 kilos the force  $x$  would be about 15, while  $y$  is greater still. The same, or even a greater difference, occurs in the arm, the tension of the muscles and the pressure in the bone being much greater than the weight supported.

Many problems in mechanics, especially with regard to moving mechanisms, may be most conveniently solved by a direct application of the law of conservation of energy, in a form known in the classical treatises on dynamics as the 'principle of virtual velocities.' If one end of a machine be caused to travel through a small distance  $p$  under the action of a force  $P$ , acting in the direction in which the motion takes place, then the force does work on the machine of amount  $Pp$ ; if the machine is a perfect one, without friction of any kind, it must give out the same amount of work at the other or driven end: suppose this end moves through the distance  $q$  while the former (driving) end moves through  $p$ , then the force  $Q$  exerted at that end in the direction of motion must be such that  $Qq$  the work done  $= Pp$ .

Hence  $Q = \frac{Pp}{q}$  and the force exerted at the driven end can be

determined by means of the geometrical construction of the machine. If there is no friction, the reasoning just given holds, to whichever end the force is applied: the machine is reversible; and if worked backwards,  $Q$  being a known applied force,



$P = \frac{Qq}{p}$ . But if there is friction, the work given out will be less than the work put in : hence if the end  $p$  is driving the other,  $Pp > Qq$ , so that  $Q < \frac{Pp}{q}$ ; whilst if  $q$  is the driving end (the machine worked backwards),  $Qq > Pp$ , and  $P < \frac{Qq}{p}$ . Thus in any case friction diminishes the force exerted by the machine.

As an example, let us take a screw press, such as a copying press ; and apply a force equal to 10 kilos-weight on the handle, the distance of which from the axis of the screw is 15 cms. ; and suppose the pitch of the screw (i. e. the distance it moves forward in one revolution) to be  $\frac{1}{2}$  cm. Then if we choose for  $p$  the movement of the driving during one complete revolution of the handle, this is  $2\pi \times 15 = 94$  cms., while the corresponding movement  $q$  of the driven end is  $\frac{1}{2}$  cm., consequently  $\frac{p}{q} = 94 \div \frac{1}{2} = 188$ , and the pressure  $Q$  exerted at the driven end would be 1,880 kilos, if there were no friction. Actually the friction is large and reduces this pressure greatly ; but still the magnification of the force applied is very great. The friction in an ordinary screw is usually so great that the machine will not work backwards at all : this is a valuable property, since it prevents the screw from running back when the driving force is removed.

The action of machines, as well as of the mechanical contrivances of the animal body, can be discussed by the aid of the method of compounding and resolving forces already described, except in one case, that in which the forces to be dealt with are in parallel lines. When that is so, the lines representing them on a diagram cannot be made to form the sides of a triangle, and consequently some other way of finding their joint effect is required. This is most conveniently done by means of the *moment* of the forces : if a body be fixed at one point (say the beam of a balance, at its central support), a force applied elsewhere will in general have a tendency to turn the body round the fixed point, and its tendency to do so, or moment, will be greater, the further off the line of the force is from the fixed point (e.g. the further a weight is put from the centre of the beam, the more it tends to turn the beam round). The exact measure of the moment of a force about a point is the product of the force into the perpendicular distance from the point to the line of the force : e.g. Fig. 14 the force  $F$

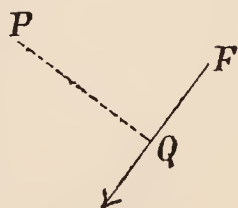


FIG. 14.

exerts round  $P$  a tendency to turn measured by the product of  $F$  into the length  $PQ$ . That this is so may be verified by considering the work that

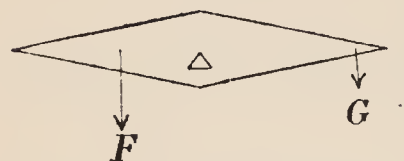


FIG. 15.

a force might do, i.e. by the principle of virtual velocities: thus supposing two forces  $F$  and  $G$  (Fig. 15) to act on the beam of a balance,  $G$  being twice as far from the centre as  $F$ ; then  $F$  tends to turn the beam against the hands of a clock,  $G$  in the clockwise sense. Suppose the beam to turn a little, say clockwise, then it is clear that  $F$  and  $G$  will both describe parts of

circles round the centre of the beam, and  $G$  move twice as far as  $F$ . If then  $F$  be twice as great a force as  $G$ , the work done (force  $\times$  distance) by  $G$  will be equal to that done against  $F$ : and this is the condition for equilibrium, since if  $G$  gave out more energy than was absorbed by  $F$  the difference would go to give velocity (and kinetic energy) to the system, whilst if  $G$  gave out less energy than  $F$  absorbs the movement would take place in the opposite sense. Hence it is only when the work done on either side is the same that the two forces can balance and leave the beam at rest: and this is the case when the moments of the forces are equal. The moment of a force actually applied to turn a shaft round is usually called the *torque* on the shaft.

The work done by a rotating shaft can be most conveniently expressed by means of the torque exerted in it. For, let  $n$  be the number of revolutions per second, and suppose the shaft to be driven by a force  $F$  applied at a radius  $r$  from the axis of rotation. Now in one revolution the travel of the point of application of the force is  $2\pi r$ ; hence in one second  $2\pi rn$ . The work done per second is therefore  $2\pi rnF$ . But  $rF$  measures the torque applied; hence we may write

$$\text{work done} = 2\pi \times \text{no. of revolutions} \times \text{torque}.$$

The arrangement just considered is a lever, and the principle of moments has sometimes been known as the principle of the lever. It can be applied to determine the force that can be exerted by such a mechanism; thus if in a pair of nut-crackers (two levers hinged together)

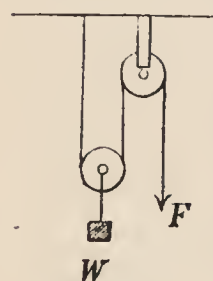


FIG. 16.

the hand exerts a pressure of 20 kilos at 12 cms. from the hinge, and a nut be placed 2.5 cms. from it, the moment of the pressure is  $20 \times 12 = 240$ ; this must be balanced by the moment of the resistance of the nut; if the latter is  $x$  kilos the moment is  $x \times 2.5 = 240$ , whence  $x = 96$ .

This, it will be seen, is practically the same as applying the principle of virtual velocities direct. Another instance of this general principle may be found in a system of pulleys; thus in Fig. 16 with one movable and one fixed pulley it may be seen from the geometry of the system that the force  $F$  applied to the rope must move 2 cms. for each 1 cm. that the



weight  $w$  is raised: hence, neglecting friction, the force need only be half the weight. The fixed pulley makes no difference to the magnitude of the force, serving only to change its direction.

Since weight always acts downward, i.e. in parallel lines, the resultant of any number of weights, or of the weight of different parts of a body, can be found by the principle of moments. If a single solid body be considered, we must take the weight of each small part separately and compound them together, one after another, till we have the resultant of the whole; in this way may be found the line along which the resultant or total weight acts. It may be shown that if the body be put successively in any number of different positions the line of the resultant weight will always pass through one point: this is called the *centre of gravity* of the body. If then a body is to be supported, the supporting force must be applied vertically under or over the centre of gravity: e.g. if a string be tied to a chair and the chair lifted, it will not in general produce equilibrium, but the chair will swing about, and eventually settle in such a position that its centre of gravity is vertically under the point of attachment of the string. Again, if a body is resting on a base, the vertical through the centre of gravity must fall within that base; this is the case in a chair as it usually stands, the base being, here, the quadrilateral formed by the four feet. But if the chair be tipped back till the vertical through the centre of gravity comes to lie outside that quadrilateral, the chair will fall over.

## § 7. Oscillations and Waves.

The most familiar instance of an oscillatory or vibratory motion is the common pendulum. To study the motion imagine such a pendulum constructed by hanging up a small ball of lead by a long thread, and set swinging, but only to a small extent, in order that we may avoid complications as much as possible. The ball then swings to and fro, in what is really an arc of a circle, drawn round the point of attachment of the thread, but it is so short an arc as to be nearly straight, and we may treat it as such. Then we may note these points about the motion; first there is a position of rest for the pendulum, viz. when it is hanging straight down; next when pulled aside from that position, and let go it tends to go back to its position of rest. This second fact shows that the position of rest is a *stable* one. The use of this term involves a knowledge of the different kinds of equilibrium; a body is said to be in equilibrium when the various forces applied to it balance one another, and so do not cause it to move: that is the case e.g. when an egg is

laid on its side, for the resistance of the table acts in the same vertical line as the weight and balances it ; but it is also the case if the egg be very carefully placed on its end. The resistance of the table is again in the same vertical line as the weight, and balances it, so that the egg will remain at rest. The difference between the two cases is that if some slight accidental disturbance moves the egg out of place when laid sideways it will return to its former position : it is then said to be in *stable equilibrium* ; but, if placed endways, the least disturbance will make it fall over : it is then said to be in *unstable equilibrium*. When it is neither on end nor on its side, but in an intermediate position, it is not in equilibrium at all. There is a third case of equilibrium, called *neutral*, which is exemplified by a ball placed on a horizontal table : the ball is obviously at rest, and, moreover, if moved slightly, it neither tends to fall back into its former position nor to fall away from it, but remains in any position in which it is put. Now if a body be in neutral or unstable equilibrium, and be slightly displaced, it will not tend to rock about that position, since it has no tendency to return : an oscillation, then, is always a *movement round a position of stable equilibrium*.

Returning to the observation of our pendulum we see that, when pulled aside, the ball is slightly raised, and consequently possesses a small amount of potential energy (work stored up in it on lifting) as compared with its position of rest. When the pendulum is let go, this potential energy is gradually converted into kinetic, as the ball acquires more and more velocity, until when the position of rest is reached, and the ball is as low as possible, there is evidently no potential energy left ; and we know, by the law of conservation, that just an equal amount of kinetic must have been generated. The ball will consequently not stop here, for that would mean that a certain amount of energy present in it vanished, which of course is impossible. Accordingly the ball goes past its position of equilibrium, and only stops when it has risen to an equal height on the other side, when all its energy is again in the potential form. This process is repeated an indefinite number of times, and constitutes the vibration : the only limit to it lies in the fact that in any real pendulum there is a small amount of friction (against the air, and at the place of support), so that at each swing a little energy is used up, and the store becomes smaller and smaller, the vibration gradually dying away.

In a vibration of the kind just mentioned the most important quantities to record in describing it are (i) the *amplitude*, by which is meant the extreme distance the vibrating body reaches on either side of the position



of rest, (ii) the *periodic time* of vibration, i. e. the time taken to swing completely to and fro, since after that the motion is merely repeated. If a common pendulum be observed with a watch and telescope, it will be found that, although the vibrations gradually diminish in amplitude, the time taken to execute one swing to and fro remains constant: the smaller vibrations are carried out with smaller velocity. This fact is observed with all vibrations of similar kind, so that any oscillation is characterized by a definite periodic time. The reciprocal of the periodic time, i.e. the number of vibrations executed in one second, is called the *frequency*.

Such an oscillation may be regarded as the projection of a uniform motion in a circle: indeed it is easy to see how it might be derived from circular motion. Suppose the ball of the pendulum in the illustration to be pulled aside and then started moving in a small horizontal circle: it will be found that it travels round and round with uniform speed, indefinitely, except that friction slowly brings it to rest as in the previous experiment: in this case it never passes through the position of rest at all in its motion, but behaves like a planet circling round and round the sun. Now go a good distance off, and, keeping the eye on a level with the ball, watch the movement; we then obtain a 'projection,' for we are able to see the movement to right and left of the eye, but not the movement towards and away from it: and it will be found that all the characteristics of the usual pendular movement are reproduced.

We may use this as a definition of the kind of movement we are concerned with. Any motion that repeats itself at regular intervals is called harmonic; and the movement of the ordinary pendulum is of the kind called simple harmonic. *A simple harmonic motion (S. H. M.) is the projection on a diameter of uniform motion in a circle.*

Numerous other instances of simple harmonic motions occur, especially when the movement is due to elastic forces. A solid, when in a position of stable equilibrium, may be subjected to a small strain and then released, and it will oscillate to and fro about the equilibrium position until friction causes it to come to rest. To this class belong the motions of tuning-forks, spring time-markers, and the vibrating parts of musical instruments: we have already referred to the tuning-fork in connexion with the chronograph (p. 4), the instrument there described, however, being provided with an artificial arrangement for maintaining the vibrations against friction indefinitely. We will here, therefore, choose as an example of harmonic motion executed under elastic forces, the time-marker shown in Fig. 3. In this a flat band of steel  $s$ , which would normally be horizontal, is bent out of that position by the weight  $w$ : it consequently attains a new position of equilibrium, under the joint action of the weight and the elastic reaction set up in it—a position such as that shown in the figure. If now it be depressed a little further the elastic reaction over-

balances the weight, and on releasing  $s$  it will spring upwards under the action of the resultant force : but when it has recovered its position of rest it has acquired a considerable velocity, and will, of course, not stop suddenly, but rise further ; and as then the spring will be less strained, the elastic stress in it will be diminished, and will no longer balance the weight, and the resultant force will then be downwards. In this way the motion repeats itself at uniform intervals. The example, moreover, enables us to see what that most important quantity, the periodic time, depends on. The weight  $w$  can be shifted to any desired position on the spring, and, of course, the further out it is the wider arc it has to swing through, and so for any given frequency of vibration of the spring the more rapidly it moves : but, as nearly all the mass of the moving body lies in  $w$ , this means that when the weight is further out the kinetic energy corresponding to any given time and amplitude of swing will be greater. However the kinetic energy can only be the equivalent of the work done by the spring in recovering its equilibrium position after being bent : if it be bent to a fixed extent, then in recovering it will do a fixed amount of work, and communicate a fixed velocity to  $w$ , so that when the weight is put further out the vibration must be executed more slowly. On the other hand, if the spring were made stronger, the weight remaining fixed in magnitude and position, then for a given amount of bending more work will be stored up in the spring, and as this work again spends itself in giving kinetic energy to the weight, the latter will acquire a higher velocity, and the vibration be exerted more quickly. Thus we see that anything that increases the capacity of the system for kinetic energy makes the period longer, anything that increases its capacity for potential energy, when displaced from the position of equilibrium, shortens the period.

In the practically important case of the common pendulum, if it be constructed of a small heavy ball suspended by a thread, it may be shown

that the time of vibration or period  $T = 2\pi \sqrt{\frac{l}{g}}$ , where  $l$  is the length from the suspension to the centre of the ball, and  $g$  is the acceleration of gravity. This is the method actually used for measuring  $g$ .

In addition to the amplitude and period, another piece of information is needed to determine the position of a vibrating body at any time. The body is said to pass through various *phases* in the course of its vibration : the term is one well known in connexion with the periodic changes in appearance of the moon—new moon, full moon, last quarter, &c. : and it is equally applicable to the set of changes constituting any other periodic phenomenon. To express phase quantitatively it may be defined by means of fractions of the whole period : thus, reverting to the case of the pendulum, if we call its phase 0 at the moment when it passes through the centre towards the right, then when it has reached its greatest elongation



towards the right its phase is  $\frac{1}{4}$ , when passing through the centre towards the left  $\frac{1}{2}$ , when at the furthest point towards the left  $\frac{3}{4}$ , and when again passing through the centre towards the right 1 or 0 again, since then the motion repeats. Now clearly if we know the phase at some moment of time, and also the period of vibration, we shall be able to calculate the phase at any future time, and, with the aid of that and the amplitude, shall know the exact position of the vibrating body at that moment.

When a vibrating reed or tuning-fork is used in connexion with the chronograph, a wavy curve is obtained such as that shown in Fig. 8; this consists of short pieces similar to each other, each leading into the next: each piece corresponding to one vibration of the reed or fork. The piece of curve corresponding to a single vibration may be of other shapes than that shown in Fig. 8, without interfering with the regular repetition, as may be seen in sphygmographic records. These other, and usually more complicated shapes, correspond to a more complex harmonic motion on the part of the vibrating body: but a reed or tuning-fork vibrates in a manner that is very nearly simple harmonic, and the corresponding tracing on the chronograph is of the shape known as a *sine-curve*. The sine-curve, then, is the resultant of two motions, one a S. H. M. in one direction (usually vertical), imparted by the tracing-point of the fork, the other a uniform motion in a straight line at right angles to the former direction, and imparted by the drum of the chronograph.

Such a tracing may serve as an illustration of a *wave*, for a wave consists in a harmonic motion propagated in some direction, usually with uniform velocity. As a simple instance of the production of a wave the following experiment may be considered: tie one end of a string to a fixed point on the far side of a room, and holding the other in the hand, give it a vibratory motion; or better, in place of string, take a long piece of rubber tubing, weighted by filling it with water. With this apparatus we may observe that by giving a single to and fro movement to the end, a disturbance is produced, which travels down to the other (and even back again by reflection); if only a single vibration is executed, this transmitted to a distance constitutes what is called a *pulse*: if the vibration be kept up, one pulse after another is transmitted, till the whole string is set in vibratory motion, and a regular set of waves is formed. The leading points to be observed—they may most easily be noticed in the case of a single pulse—are (1) each point of the string executes a movement similar to that of the starting-point, (2) the movement occurs later and later as the distance from the starting point increases: i.e. the moment at which each point passes through a given phase is later according to the distance of the point from the starting-point. By observing the time when successive points reach the same phase, we may measure the velocity with which the wave is propagated; this will be found uniform in the case chosen for experiment, and in



general ; in fact, the *velocity of a wave usually depends only on the nature of the medium in which the wave is transmitted.*

When a complete wave is set up, the movement of any individual pulse may be watched, and the velocity of the waves determined by means of it. But by the time that one point A (figure, p. 246) has reached a particular phase—say its greatest elongation to the right—another point B, nearer the place of starting, will be found to be in the same phase, being really one whole period in advance of A. The distance AB is called the *wave-length* ; it is, therefore, the distance that the wave travels forward during the time of one complete vibration, and in the tracing representing the wave (in the figure, p. 246, a curve of sines) it is the length after which the curve repeats itself. It follows then, from this definition, that

$$\text{wave-length} = \text{velocity of wave} \times \text{periodic time} ;$$

or since frequency means the reciprocal of the periodic time, we may write this—

$$\text{velocity of wave} = \text{wave-length} \times \text{frequency}.$$

The observations made on the vibrating cord may also be made on the ripples set up in a pond by dropping a stone in it, with this difference however, that as the waves spread out into larger and larger rings, the amplitude of vibration set up in the water falls off.

This is because the wave only possesses a fixed amount of energy, and as it spreads out it sets more and more water in motion, and consequently cannot impart so large a vibration to each point of the water. The same is true of waves—such as those of sound—which spread out in all directions in space ; but not when the waves are propagated in one direction only, as in the experiment with the string. Here there is nothing to make the amplitude fall off except the gradual consumption of energy in overcoming friction.

## § 8. Hydrostatics.

A vessel containing a liquid suffers everywhere a pressure on its walls, tending outwards ; as may easily be shown by making a hole in the wall, when the liquid rushes out in consequence of the pressure. To find the amount of this pressure, suppose a vertical tube, 1 sq. cm. in cross section, to be filled to a height  $h$  with a liquid of density  $d$  : then the tube contains  $h$  cubic centimetres of liquid, or  $hd$  grams. The weight of this has to be borne by the base, which consequently suffers a pressure of  $hd$  grams weight per square centimetre ; but adopting, as before, the systematic unit of force, this amounts to  $hdg$  dynes per sq. cm. ( $g = 981 =$  the acceleration of gravity.)

Observation shows that however a vessel, or set of communicating vessels, be shaped, the liquid will stand at the same level throughout it. This means that the pressure at any given depth below the surface must be the same however the vessel may be shaped: if, for instance, the vessel cone outwards from the bottom up, the weight of water it contains is greater than that standing vertically over the base; but the excess is borne by the sloping sides, leaving the pressure still  $= h\delta g$  on each unit area of the base. So again if it cone inwards the pressure remains the same, although it is now greater than the whole weight of water contained: the difference being accounted for by an upward pressure on the sloping sides. Not only is this the case, but there is an equal pressure on the sides of the vessel whatever direction they have: hence, e. g., if a closed vessel (such as a thermometer bulb) have a tube filled with liquid rising from it, the pressure on its sides may be very considerable, depending as it does on the depth below the free surface of the liquid, and not merely on the weight of liquid contained in the tube.

The pressure of the liquid on the base of the vessel must not be confused with the pressure that the vessel exerts on the support bearing it. The latter is, in any case, equal to the whole weight. Thus, taking the example of a vessel coned outwards, some of the weight of water is borne by the sloping sides, and some by the base, but the whole ultimately falls on the table on which the vessel stands.

Gases have weight, and consequently produce a pressure in precisely the same way as liquids: hence the great quantity of atmospheric air overhead causes quite a large pressure on any surface exposed to it (including that of the human body, which, however, is so accustomed to the pressure as not to notice it). This can easily be shown by the instrument known as a *barometer*: if a glass tube some 90 cm. long, ending in a tap, be placed with the tap uppermost, and a reservoir of mercury be attached by flexible tubing to the lower end, then by opening the tap and raising the reservoir, we may drive out the air through the tap and fill the tube completely with mercury; then if the tap be closed, when the reservoir is lowered the mercury in the tube will no longer be pressed on by air, whilst that in the reservoir is; and it will be found that the free surface-level



in the tube is higher than in the reservoir by some 76 cm. Hence the pressure due to the weight of air overhead balances that due to a column of mercury of the height stated. The height varies somewhat from day to day, and from place to place, and a barometer is intended chiefly to measure it; but as a standard value, 76 cms. of mercury at the temperature of  $0^{\circ}$  is chosen. Now the density of mercury at that temperature is 13.596, and the mean value of  $g$  is 980.61. Hence the normal atmospheric pressure is  $76 \times 13.596 = 1033.3$  gms. weight per sq. cm.

or  $76 \times 13.596 \times 980.61 = 1,013,200 \frac{\text{dynes}}{\text{sq. cm.}}$ . Sometimes, however,

$1,000,000 \frac{\text{dynes}}{\text{sq. cm.}}$  is chosen as the standard pressure.

Gauges to measure fluid pressure are usually made of a column of liquid in a glass tube. The most familiar is the barometer, already mentioned: it measures the actual pressure of the air, because one end of the mercury (that in the reservoir) is pressed on by the air, the other end (under the tap) by nothing, so that the whole pressure of this air has to be balanced by that due to the difference in level of the two mercury surfaces. Barometers are not usually made with a tap and flexible tube, as described above, but either of a simple glass tube, bent into a U shape, with one end sealed up, the other open to the air; or of a straight glass tube, sealed at the top, and dipping into a dish full of mercury. Mercury is chosen because (1) it is so heavy that the barometer is shorter than when made with any other liquid; e.g. a glycerine barometer would need to be about 8 metres high to produce the pressure sufficient to balance the air; (2) it is almost involatile. If a volatile liquid like water were used the space at the top of the tube would be filled with vapour, and this would produce a considerable downward pressure on the liquid, and the height of the liquid column would only represent the difference between the air-pressure and that of water vapour (see p. 96). The chief precautions required in making a barometer are: (1) that the mercury should be pure (preferably distilled *in vacuo*); (2) that the tube be completely free from air-bubbles and moisture. To get rid of moisture the tube must be heated strongly: a simple and convenient form for the purpose is that represented in Fig. 17, with an air-trap A: the tube can be filled by merely pouring in mercury; this is done



partially, and the portion above the trap then freed from air and moisture by boiling repeatedly over a bunsen burner; it is then completely filled with mercury, the finger put on the open end, and the tube inverted into a dish of mercury. If any air-bubbles remain in the lower part of the tube they collect under the trap, where they do not affect the vacuum at the top. The tube should not be less than 8–10 mm. wide at the top, to avoid capillary effects (p. 112). The most important correction to the readings of the barometer is for temperature: as mercury expands with heating, it is desirable to correct the height of the column to that which it would assume if the mercury were at the standard temperature of the freezing-point: to do this, if a glass scale is used, subtract 0.13 mm. for each degree above zero; if a brass scale, 0.124 mm.

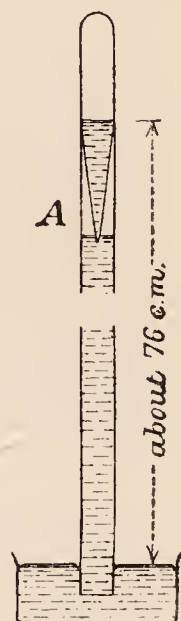


FIG. 17.

When a gauge is only intended to measure the difference between two pressures, the construction of it is easier; it may be necessary, e. g., to compare the pressure of the gas in a piece of apparatus with that of the air outside, and this can be done with a simple U-tube of mercury. When gas is enclosed in a vessel, it produces on the walls of the vessel a hydrostatic pressure comparable with that existing in the atmosphere on account of its weight. Thus, suppose we have a glass bulb fitted with a tap, and let the tap be open, then obviously the atmospheric pressure, 760 mm. of mercury, prevails throughout. Now let the tap be closed: no difference has been made to the air inside, so that a hydrostatic pressure of 760 mm. still exists in it, although it is no longer directly pressed on by the weight of air overhead. This fact is explained by the view that a gas consists of small particles (molecules) which are in rapid movement in all directions, and so very frequently collide with the walls of the containing vessel, and tend to drive them outwards. It is this pressure, due to the velocity of the molecules, which balances the applied pressure the gas has to support, whether that be due to the weight of more gas, or be applied by means of a mercury or other liquid column, or be due to the resistance to expansion of the solid containing vessel. Moreover it is observed that the pressure produced in a given vessel is proportional to the mass of gas it contains (Boyle's law, see p. 77);

so that if twice as much air be forced by means of a pump into the bulb in the above experiment, with the tap closed, the pressure inside it would be sufficient to balance  $2 \times 760 = 1520$  mm. of mercury.

Hence we are justified in speaking of the pressure inside a vessel of gas, as well as that in free air: and whilst in liquids it is usually necessary to bear in mind that the pressure is different at different depths, owing to the weight of the liquid, this is hardly ever necessary in the case of a gas, because its weight is so small. Thus if an apparatus containing air were five metres high, the pressure at the bottom of it would be greater than that at the top only by

$$hdg = 500 \times 0.0012 \times 981 = 58.86 \frac{\text{dynes}}{\text{sq. cm.}}$$

or about  $\frac{1}{20000}$  atmosphere.

To measure the difference between the pressure in two vessels containing gas, or between one vessel and the atmosphere, a simple U-tube of mercury is sufficient, each end of the U being connected to one of the vessels. If the pressure difference is small, a lighter liquid than mercury may be used with advantage, such as sulphuric acid, or oil; and since these liquids wet glass, no capillary error will be made, provided the two limbs of the gauge are of the same bore, consequently the gauge may be made much narrower: 1 or 2 mm. diameter is sufficient. For registering pressure on the chronograph (i.e. differences of pressure between some apparatus and the air), a float in the open end of the gauge is made to carry the tracing-point. It should be noted that if the gauge is to respond rapidly to changes of pressure, such as those occurring in the vascular system, its moving parts must be made light, for the reasons already discussed in the case of levers (p. 16). But if the mean pressure only is desired a mercury gauge is satisfactory because its considerable inertia prevents its oscillating too easily.

Other methods have been adopted in order to avoid, more completely, the disadvantages of inertia in a pressure gauge; especially the plan of balancing the pressure against the elasticity of a spring of some kind. Thus Fick's 'kymograph,' or manometer, consists of a C-shaped spring, made by uniting two thin strips of metal at the edges, so as to form a flat tube: if pressure be applied to the air in the interior of the tube, the C opens out and so registers the pressure. Aneroid barometers are con-



structed on the same principle, the internal space of the tube being evacuated, so that the tube is more or less coiled up according to the pressure of the air acting on its exterior surface. Again, Hürthle's manometer consists essentially of a drum, or tambour, covered with thick india-rubber; the interior of the drum is filled with salt solution, which transmits the pressure from the blood-vessels, which it is desired to register: when the pressure increases, it raises the drum-skin, which by a short metal arm moves a very light lever, provided with a tracing-point to record the movements on the chronograph.

One of the effects of fluid pressure is that a solid immersed in a fluid is pressed upwards, and consequently loses part of its weight. The upward pressure amounts to the weight of the fluid displaced. The effect is most obvious in liquids on account of their possessing a much greater density than gases, but it occurs also in the latter. Two cases have to be distinguished, according as the solid is denser or less dense than the liquid in which it is immersed. In the latter case the solid floats, a portion of it projecting above the liquid surface; that is to say, just enough of it is immersed to make the weight of displaced liquid equal to the weight of the solid itself, and the solid is thus practically relieved of weight altogether. Such a case has already been mentioned in the common hydrometer (p. 11). If the solid is denser than the liquid, it sinks completely below the surface: it has then displaced a bulk of liquid equal to its own, and is apparently lighter by the weight of that displaced liquid, as may be proved by attaching it, while immersed, to the arm of a balance (p. 12). In the intermediate case, in which the solid and liquid are of precisely the same density, the solid will not necessarily settle either at the top or bottom of the liquid, but will remain freely anywhere, having lost all its weight. This case is sometimes observed in the use of specific gravity heads (p. 11).

But if a solid lighter than water be forcibly immersed below the level at which it naturally floats, its apparent weight becomes negative, i. e. the upthrust of the liquid more than balances its weight. A boat, e. g., floats freely at a certain level: if a man gets into it, it floats deeper, and the additional water so displaced produces an upward pressure just sufficient to balance the weight of the man, which is the external force pressing the boat downwards.



The hydrostatic support of an immersed solid may be simply proved by the energy principle. For suppose a small piece of solid to be immersed to a depth  $h$  below the surface of a liquid, and let  $m$  be the mass of liquid displaced by it: then if the solid were caused to rise to the surface, the liquid, flowing down to take its place, coming as it would practically from the surface, would do an amount of work = weight ( $mg$ )  $\times$  height =  $mgh$ : this amount of work therefore would contribute towards raising the solid, and the force exerted on the solid is obtained by dividing the work done by the distance through which the force is exerted, or  $mgh \div h = mg$ , the weight of the displaced liquid. And this result is clearly true for every small piece of a solid of whatever size or shape, and hence though the large solid may not all be at one depth below the surface, the result is still true.

All the bodies we are familiar with are immersed in an ocean of air, and hence lose weight, although very little. This fact must be taken into account in very accurate weighings. Also, though no actual solid is so light as air, an arrangement may be constructed of a light bag filled with hydrogen or coal gas (balloon) which weighs less than an equal volume of air, and is consequently pressed upwards in air, like a piece of wood in water.

One of the most important practical problems in connexion with the pressure of gases is that of increasing or decreasing it, i.e. the construction of *pumps*. Mechanical pumps for compression or rarefaction depend essentially on the working

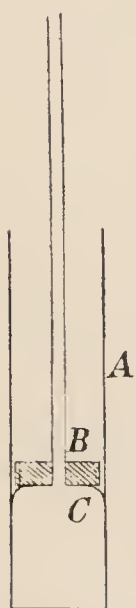


FIG. 18.

of a piston in a cylinder provided with valves. The ordinary cycle-pump may be taken as an example: it consists of a cylinder  $A$ , in which slides a piston  $B$ ; the latter carries a flange  $C$  of leather or india-rubber attached somewhat aslant, so that when the cylinder is pulled outwards, making the volume under the piston greater, air flows in past the flange; but on reversing the stroke the air in attempting to get out again presses the leather against the wall of the cylinder, and closes the aperture: the flange consequently serves as a valve in this apparatus, allowing the air to flow one way, but not the reverse. The air, thus compressed in the space between the cylinder and piston, escapes through the piston-rod, which is hollow. The tube by which this compressed air is delivered into the vessel intended to contain it (e.g. the cycle-tire) must have another valve, to prevent its flowing back

during the next outstroke of the pump. A similar arrangement with the valves reversed would serve for rarefaction; but when the vacuum was very good the air-pressure would no longer be sufficient to work the valves: hence to get the best results. they should be worked mechanically, as for instance in the Fleuss pump (of which the essential parts are shown in Fig. 19): it consists of a cylinder *M*, in which a piston *N* works: the piston-rod carries the valve *G*, opening upwards. Both the piston and the valve *G* are completely covered by a non-volatile oil. The piston on its up-stroke passing the point *B* leading from the surrounding brass chamber, forces the air in the cylinder up through the valve *G*, and the oil following the air leaves no waste space: on the down-stroke the oil serves to close the valve *G* perfectly: a vacuum is consequently produced over the piston, and air flows from any receptacle connected to *A* through *B* into the cylinder, and the stroke is repeated. No air-pressure is needed to work the valve *G*, as it is lifted by the collar *O*, and replaced by a spring. It

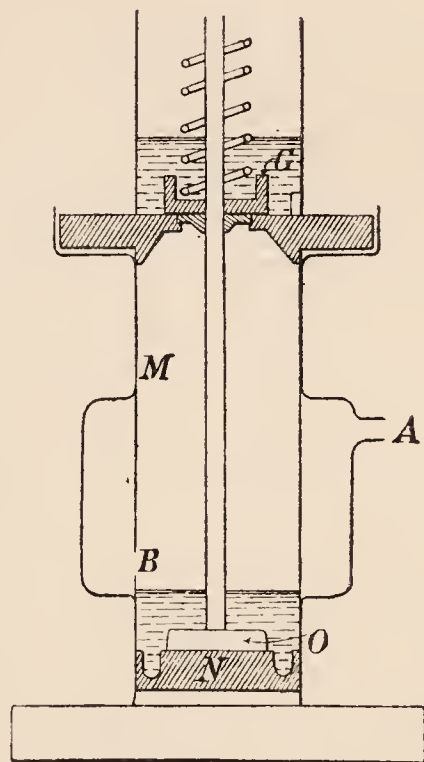


FIG. 19.

is possible to arrange two such pumps 'in series,' i.e. to make the second pump the air out from the barrel of the first. By this means the air-pressure can be reduced to a small fraction of a millimetre, and a vacuum obtained good enough for the manufacture of incandescent lamps and 'vacuum-tubes.'

Most mechanical air-pumps are not however capable of reducing the air-pressure below one or two mm.; for producing much better vacua than this mercury-pumps are mostly used. These may consist of a cylinder with a piston of mercury, instead of a solid: the Töpler pump shown in Fig. 20 is actuated by lifting the reservoir *A* of mercury up and down. As it rises, the mercury fills the cylinder *B* and forces the air out by the capillary tube *c*; when the stroke is reversed, the pressure of the atmosphere drives mercury up *c* (which must be at least 76 cms. long) and so prevents return of the air; hence air flows

into B from the apparatus to be exhausted, through D, which is usually provided with a drying-tube containing phosphorus pentoxide. On the next stroke the air in the cylinder cannot

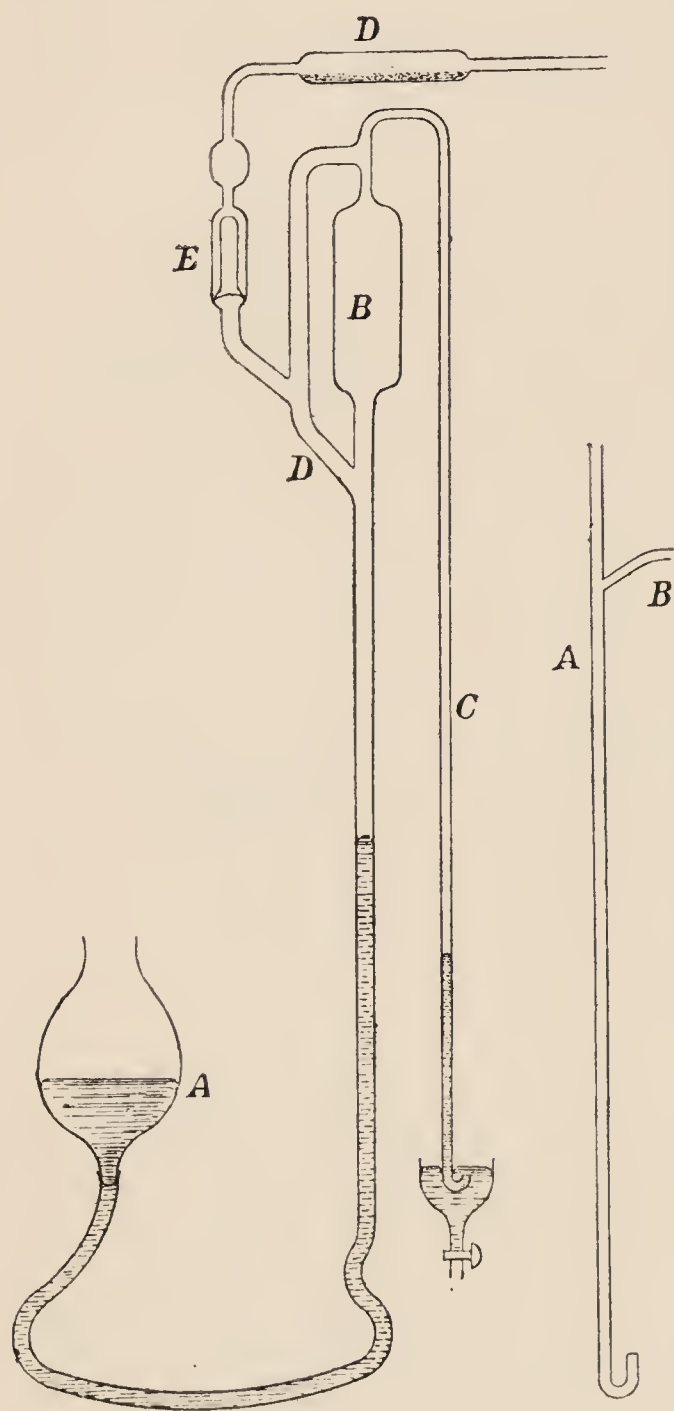


FIG. 20.



FIG. 21.

get back into the vessel to be exhausted, because the mercury closes the tube D: the valve E is to prevent mercury being driven up too high into D; when A is lowered, it falls by its own weight and leaves the passage free. By means of a mercury-pump the air can be all but completely removed, the gas remaining in the apparatus being mainly mercury vapour, which has a pressure at ordinary temperatures of about  $\frac{1}{50}$  mm. Mercury-pumps are now made which are actuated mechanically by means of water-pressure, so that it is only necessary to turn a tap and leave the exhaustion to proceed automatically: such have been designed by Raps, and by Kahlbaum.

Another mercury-pump, acting on an entirely different principle, is that of Sprengel (Fig. 21). In this the mercury falls in drops down the tube A, and in

passing the mouth of the side-tube B each drop carries a small quantity of air with it. The tube A should be of about 1.5 mm. diameter, and long enough for the mercury in it to balance the atmospheric pressure.



A similar pump can be worked with water (Fig. 22). Here drops falling from the capillary tube B drag down with them air that flows into the chamber A through the side tube C. The tube D that leads away the water should be long; if it be only a metre or two, the pump will only diminish the pressure a little, and is chiefly suitable for drawing a slow stream of air through any apparatus, but by using a fall tube of 10 metres a good vacuum can be got. If C be left open, the air delivered at the bottom of D will be under a small excess of pressure and may be led away to an apparatus through which it is desired to drive a slow stream of air, e. g. for artificial respiration.

Another water-pump is that shown in Fig. 23, the theory of which is explained below (p. 117).

The flow of water through the fine jet A sucks air through the side-tube B. The action of this pump is much more rapid than that of the water-Sprengel, and if the water-pressure be sufficient it may be made to exhaust down to the vapour-pressure of water, i. e. about  $\frac{1}{50}$  atmosphere. The air is carried down through the outlet, and may also be separated from the water and delivered under pressure.

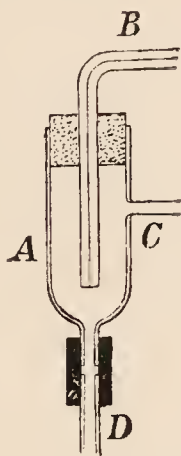


FIG. 22.

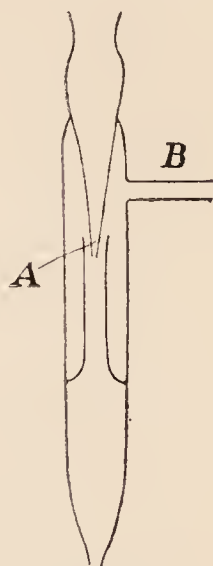


FIG. 23.

## CHAPTER II.

### HEAT.

#### § 1. Temperature.

BEFORE dealing with the various properties of matter in turn, it is necessary to consider the very far-reaching influence of temperature, for there appears to be no property except gravitation (weight) which is not affected by it. The conception of temperature is originally a physiological one, based on the sensations of hot and cold : to make it physically useful we must make the observation that hot bodies and cold ones left in contact tend to become of the same temperature ; e. g. a vessel of hot water cools when left to itself, i. e. it tends to become of the same temperature as the air with which it is in contact. Generalizing this, we may say that equality of temperature is the state in which two bodies will remain, as regards their heat, if left in contact long enough, and that one body is hotter (at a higher temperature) than another when, on putting them in contact, heat flows from the former to the latter.

Now since change of temperature affects nearly all the properties of a body, it is only necessary, in order to measure temperature, to choose some property and record its changes. Almost any property will do, and a good many have actually been used—length of a bar, volume of a liquid, pressure of a gas, electrical resistance of a wire, and so on ; if then we note that a bar of metal is of a certain length when put in a furnace, and another day is of the same length on being put into another furnace, we may conclude that the second furnace is of the same temperature as the first, and we might register a scale of

temperatures according to the length assumed by the bar in various furnaces.

Any instrument thus designed to measure temperatures is called a thermometer; and the one most in use is the mercury in glass thermometer. This consists of a small cylindrical glass bulb fused on to a glass tube, usually six or seven mm. in diameter externally, but of very small bore; the bulb and part of the stem are filled with mercury, and the top sealed, the stem above the mercury being left empty of air. When heated the glass bulb expands slightly, but the mercury expands much more, and consequently has to rise in the tube to find room: the tube is therefore provided with a graduated scale against which the position of the mercury thread may be seen. By the device of the bulb and stem even a very small change in temperature, producing a very small change in volume of the mercury, may be observed. When the instrument has been made, in order to graduate it, it is placed (i) in melting ice, (ii) in steam over boiling water. To perform the first operation successfully it is chiefly necessary that the ice should be pure: a test tube full of distilled water should be partly frozen by immersion in a mixture of ice and salt, the water in the tube being stirred to keep the ice formed in a powdery state: the thermometer is then put in the test tube and left, with occasional stirring, for a few minutes: the temperature will then be constant, and will remain so as long as there are both ice and water present in the tube: the level of the mercury is then marked. For the second test a flask with a side tube is taken, partly filled with water, and the thermometer fitted into the neck with a cork, so that its bulb comes a little way above the surface of water in the flask: the water is then boiled for a few minutes, and the level of the mercury thread noted: in the meanwhile the height of the barometer should be read. In both operations it is necessary to see that the mercury thread projects as little as possible from the ice or steam, so that the whole of the mercury may be at the desired temperature. The temperature of the ice-point is called  $0^{\circ}$  (on the centigrade scale), that of the steam-point  $100^{\circ}$  when under the normal atmospheric pressure of 760 mm.: but the latter varies with pressure, rising  $1^{\circ}$  for 26.8 mm. rise of pressure, so that the observed position of the mercury must be corrected according to the reading of the barometer. When the ice-point



and (corrected) steam-point have been found the length of stem between them is divided into 100 parts, and the graduation, if desired, is continued above and below<sup>1</sup>.

When so graduated, however, the thermometer will still not read accurately unless the stem is of uniform bore throughout: a very good thermometer must therefore be 'calibrated' to determine any small irregularities in its bore. But as this is a very tedious process, most thermometers are merely compared with a standard, by immersion in baths of various temperatures: this is done for a small fee at Kew Observatory, and at the Berlin Reichsanstalt; so that the most convenient way of measuring temperatures accurately is to purchase such a standardized thermometer. A few further points of practical importance must be noted: in the first place, two thermometers made of different materials (e. g. mercury in hard glass, mercury in soft glass), although made to agree at 0° and 100°, will usually not quite agree at other temperatures: there is no reason indeed why they should, for that would imply that the expansion of the different materials went proportionally over the whole scale of temperature. The differences on this account, though small (under  $\frac{1}{100}$ °) between 0° and 100°, may amount to several degrees at say 300°: hence it is necessary to choose some pattern of instrument as a standard and reduce the others to it. The standard thermometric scale adopted is that indicated by the pressure of hydrogen gas (or the so-called 'constant-volume hydrogen thermometer'), and in the certificates issued from Berlin and Kew the reduction to this scale is already made. Next, glass when heated and caused to expand does not at once on cooling come back to its original size, but retains a slight change, which slowly disappears: this thermal after-effect is the greatest cause of trouble in accurate thermometry. It is however less in hard glass than soft, and less in thermometers that have been kept for a long time. With a matured thermometer of hard glass it may usually be neglected.

The most important correction to be applied is when the thread of mercury is not completely immersed in the body whose temperature is to be measured. To apply this, a small auxiliary thermometer should be used to take the average temperature of the stem of the main thermometer: let  $t$  be the reading of the main thermometer,  $t'$  of the auxiliary, and  $n$  the number of degrees of the mercury thread that are exposed; then a correction

$$\frac{n(t-t')}{6400}$$

<sup>1</sup> On the Fahrenheit scale the ice-point is called 32°, the steam-point 212°; consequently a cent. degree =  $\frac{9}{5}$  Fahr. degrees, e. g. 37° cent. = 37  $\times \frac{9}{5}$  = 66.6° above the freezing-point, or 66.6 + 32 = 98.6° Fahr.

degrees should be added to the reading of the main thermometer. This is especially important in measuring high temperatures.

Thermometers are usually made either with solid stems engraved on the outside, or with hollow glass stems enclosing a scale engraved on milk-glass. The former are stronger, but are more exposed to risk of inaccurate reading due to 'parallax': this means that as the scale is some way in front of the mercury thread, the reading will appear to differ according as the eye is held high or low; for accuracy the eye must be just on a level with the top of the mercury column; this will be the case when the graduations nearest to the top appear to be in line with their reflections in the mercury. In milk-glass scale thermometers this difficulty does not arise, because the scale is almost in contact with the mercury thread. Thermometers should have a small bulb at the top, to leave room for expansion of the mercury in case it is accidentally heated above the range of the instrument, otherwise the expanding liquid will break the glass.

Ordinary mercury thermometers have, at most, a range from  $-40^{\circ}$ , the freezing-point of mercury, to  $+350^{\circ}$ , at which it boils, but it has recently been found practicable to extend the range upwards, by filling the top of the stem with a gas under high pressure: this raises the boiling-point of the mercury and allows of its use up to  $550^{\circ}$ . Such thermometers are necessarily made of hard glass in order not to give way at such high temperatures and pressures.

To make a thermometer extremely sensitive it is necessary either to make the bore very small—and there are practical limits to this—or the bulb large. Thermometers with very large bulbs have been much used of late years for Beckmann's freezing-point apparatus and other purposes. Such an instrument is shown in Fig. 24. The bulb is 10 to 12 mm. in diameter: the stem graduated in hundredths of a degree: the instrument therefore necessarily has a short range—usually some six degrees. But to extend its availability it is provided with a reversed bulb at the top: when it is desired to use the instrument for higher temperatures it is warmed until a sufficient amount of mercury has flowed into the upper bulb; then, with a shake, this is detached, and the remainder made use of in the usual way. Of course the scale

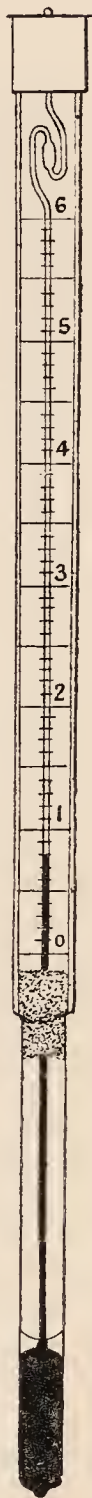


FIG. 24.



does not in such an instrument indicate any fixed temperature, but serves to measure small differences of temperature, such as that between the freezing-point of water and that of a salt solution.

For very low temperatures other liquids must be used instead of mercury. Alcohol, toluene, and petroleum ether are in use: the latter being available down to nearly  $-200^{\circ}$ .

For clinical purposes a thermometer is needed that registers the maximum temperature it has reached. This is usually accomplished by making a constriction in the tube, near the beginning of the scale: the pressure of the mercury drives it past this in expanding; but on contracting again the mercury thread remains in the tube until it is shaken back.

Partly on account of its limited range, and partly for reasons of a more theoretical kind, the mercury thermometer is not adopted as the standard, but is referred to the scale of a gas thermometer. The various gases, whether their expansion or their rise of pressure be used as indicator of temperature, give very nearly the same thermometric scale: but to be quite exact, the scale due to the increase of pressure of hydrogen, when kept at a constant volume, is chosen as the international or *normal scale* of temperature. A gas when kept under a constant pressure (as measured by a gauge, like a barometer, communicating with it) expands by 0.367 of its volume on heating from the freezing-point to the boiling-point; but if by increase of pressure it be prevented from expanding, then its pressure rises by the same fraction; so that, e. g., if a quantity of hydrogen at  $0^{\circ}$  have a pressure of 1,000 mm. of mercury, then at  $100^{\circ}$  its pressure will be 1,367 mm. nearly. Accordingly the temperature to which the hydrogen is raised may be expressed by means of the pressure it reaches.

The apparatus used is shown, a good deal simplified, in Fig. 25. A bulb A of glass or platinum-iridium contains the hydrogen, and can be raised to the desired temperature by the bath B, in which also mercury thermometers TT may be placed for comparison. The bulb is connected by a fine capillary tube C to the gauge D: this is a wide tube of glass, containing mercury. The mercury reservoir E, attached by rubber tubing, can be raised or lowered till the mercury in D is brought to a fixed mark; this ensures that the volume occupied by the hydrogen is always the



same. The other end of the gauge F is a similar wide glass tube in which the mercury can rise and fall, and which serves at the same time as a reservoir for the barometer tube G. Then since

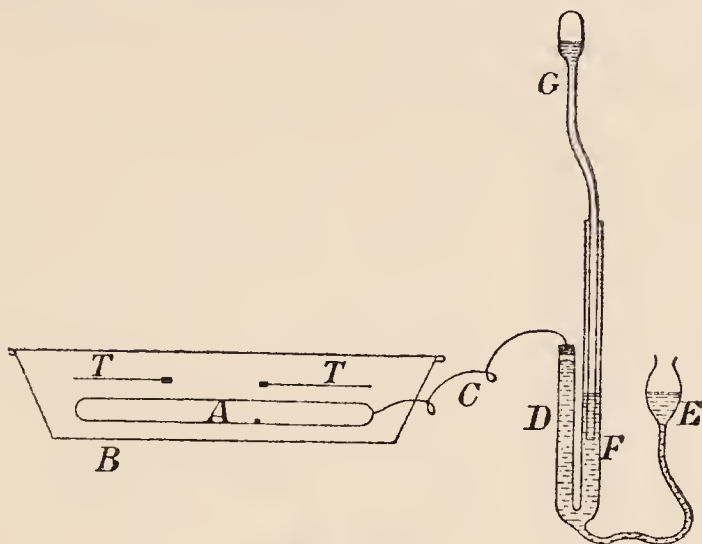


FIG. 25.

the mercury in G has a vacuum above it, and that in D is exposed to the pressure of the hydrogen, the difference in level between the two mercury surfaces measures the pressure of the gas, as explained in p. 38. The difference in level is read by means of a scale placed alongside the tubes and a telescope sliding on an upright; for convenience of reading, G is bent so as to come vertically over D: G and D must be wide tubes at the part where the mercury surface occurs, otherwise the mercury level will be irregularly affected by surface tension (see p. 112).

Since the hydrogen increases in pressure by 0.367 of its amount at  $0^{\circ}$  on heating to the boiling-point of water, and this interval is to be described as  $100^{\circ}$ , it follows that one degree may be defined by a rise of  $0.00367 = \frac{1}{273}$  in pressure: e. g. if the pressure at  $0^{\circ}$  be adjusted to 819 mm., then when on heating at constant volume it rises to 822 mm. the temperature is  $+1^{\circ}$ : when on cooling it falls to 816 mm., the temperature is  $-1^{\circ}$ . It is clear that if the temperature were reduced  $273^{\circ}$  below zero, and the same rule held, the pressure would fall to nothing. As a matter of fact the hydrogen would liquefy before that point was reached (at about  $-245^{\circ}$ ); still the supposed point of no pressure is very convenient to reckon from, and temperatures reckoned from it are called 'absolute.' *Absolute temperature, then, is equal to the temperature centigrade  $+273^{\circ}$ : it is simply proportional to the pressure in a hydrogen thermometer.* The supposed point of no pressure

( $-273^{\circ}$  cent.) is called the absolute zero. These results may be put algebraically thus: If  $p_0$  is the pressure of the hydrogen at  $0^{\circ}$  cent. (in melting ice),  $p$  the pressure at some other temperature  $t^{\circ}$  (cent.), then the absolute temperature of melting ice is  $273^{\circ}$ , the other temperature  $t + 273$ , and

$$p_0 : p = 273 : 273 + t,$$

whence  $t$  may be calculated.

In connexion with the measurement of temperature arises the equally important practical problem of the maintenance of constant temperature—an extremely important matter for many physical and chemical, as well as physiological experiments. Two methods are in use: (i) to make a bath of some material that undergoes a change at a fixed temperature, (ii) to use a bath of water, air, or other convenient material, and attach an automatic regulator to the heating supply. The changes available for use in (i) may be grouped into (a) fusion, (b) evaporation. Of the former the most familiar is the ice bath, already referred to: so long as both ice and water are present the temperature of the mixture remains constant at  $0^{\circ}$ . Many other temperatures may be obtained by means of other substances: temperatures below  $0^{\circ}$  by means of mixtures of ice and salts in the proper proportion to form ‘cryohydrates,’ e. g. NaI and ice will fall to  $-30^{\circ}$ ; high constant temperatures by the melting of metals, such as tin and lead; and by means of the chemical transformation analogous to fusion suffered by certain hydrated salts, temperatures between  $0^{\circ}$  and  $100^{\circ}$  may be got, e. g. Glauber’s salt  $\text{Na}_2\text{SO}_4 \cdot 10\text{H}_2\text{O}$  ‘melts’ in its water of crystallization at  $32^{\circ} \cdot 4$ . Secondly, the boiling-point of a pure liquid depends only on the pressure under which the boiling takes place; this has already been referred to as a means of testing the  $100^{\circ}$  point on a thermometer; other points may be obtained by boiling either carbon disulphide, alcohol, toluene, chlorobenzene, aniline, and other liquids that are easily purified and do not decompose on boiling. The liquid should be boiled in a wide glass tube or other convenient shaped vessel, inside which can be placed the apparatus whose temperature it is desired to maintain constant, and the boiling tube provided with a reflux condenser. With a few liquids only a few temperatures are obtainable in this way, but by boiling them under an adjustable pressure any temperature desired may be reached. Fig. 26 shows in outline the apparatus

required for this purpose. The barometer tube A is to be maintained at a constant temperature; it is surrounded by the jacket B, which has a small bulb c' attached in which the liquid is boiled: B is fitted at the bottom by a rubber stopper made gas-tight by a layer of mercury above it; D is the reflux condenser that keeps the

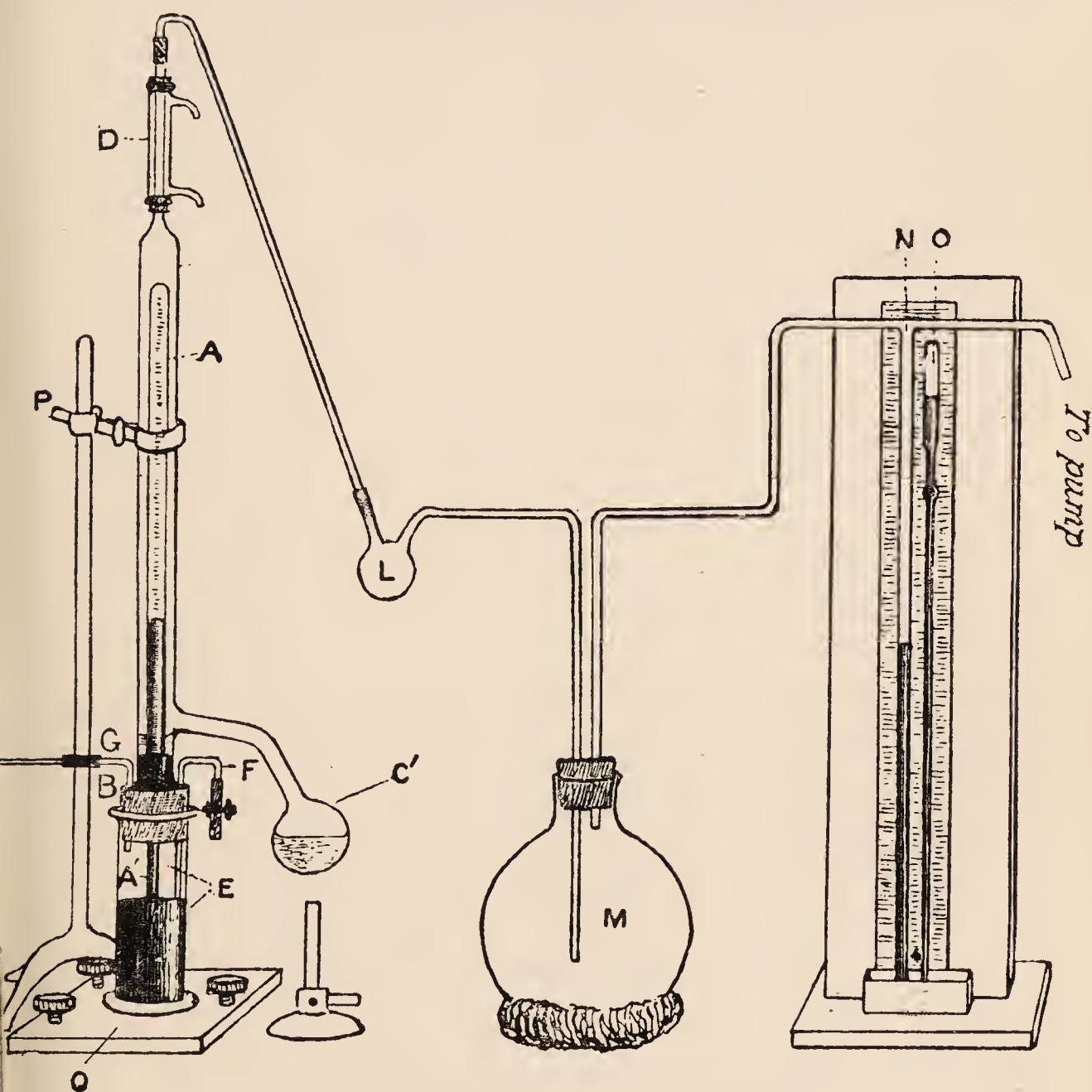


FIG. 26.

liquid in the jacket, M a large bulb to keep the pressure more steady, NO the gauge by which the pressure is measured. The appropriate liquid is placed in c', the apparatus fitted up air-tight, a burner placed under c', and the pressure reduced by the pump till the liquid in c' boils at the desired temperature. Ramsay and Young



have prepared tables showing the pressures under which certain liquids boil at each successive degree of temperature ; with the aid of these any temperature from  $0^{\circ}$  to  $350^{\circ}$  may be obtained, and when once established in the apparatus may be kept up indefinitely. The alternative method (ii) is to immerse the body

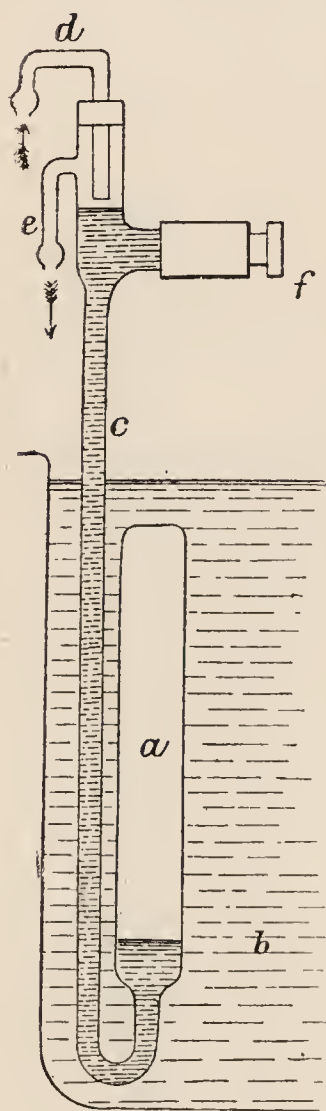


FIG. 27.

whose temperature it is desired to maintain in a bath, usually of water, or for high temperatures liquid metals, and heat the bath by a gas-burner controlled by an automatic regulator. Such regulators have been constructed depending on the vapour pressure of a liquid, but more often depending on expansion. A very common and useful type, due to Ostwald, is shown in Fig. 27. The large bulb *a* extends to the whole depth of the bath *b* and is filled with toluene or a strong solution of calcium chloride; when the temperature rises the expansion of this liquid forces mercury contained in the tube *c* upwards. In the top of this tube is fitted a glass tube *d* to carry the gas supply: the gas passes down it, over the surface of the mercury in *c* and out at *e* to the burner, but when the temperature rises to a certain point the movement of the mercury up *c* stops the gas supply: a small hole blown in the side of *d* allows just enough to flow through to keep the burner alight, so that if the temperature shows a tendency to fall and the mercury contracts away from *d* the flame springs up again: *f* is an iron screw to adjust the level of mercury and so

alter the temperature at which the regulator works. With any such device the liquid bath must be kept well stirred by means of a motor.

## § 2. Quantity of heat: dynamical equivalent of heat.

Temperature alone is not sufficient to describe the phenomena of heat. There is needed as well the quantity known as 'quantity of heat,' the quantity whose flow is regulated by differences

of temperature. The quantity of heat contained by a body is greater the higher the temperature to which it is raised, but apart from that, is proportional to the mass of the body itself, and moreover varies according to the specific character of the body. Thus a large mass of boiling water will give out heat for a longer time than a small mass of the same substance: and a kilo of hot water will give out more heat than the same mass of iron at the same temperature.

Hence to measure quantities of heat it is necessary to choose a unit in which these points are borne in mind. The unit is *the quantity of heat required to raise one gram of water through one degree in temperature*, and is called a *calorie*. Other substances differ from water in regard to power of storing heat, hence the need of the term *specific heat*; the specific heat of a substance meaning the amount of heat required to raise one gram of it through one degree; further, it is often convenient to employ the term *thermal capacity* with regard to an individual body (e. g. a flask, a calorimeter), meaning by that the amount of heat required to raise it through one degree.

Then to calculate the heat required to raise any substance through any range of temperature, it is necessary to multiply together the mass of the substance, its specific heat, and the rise of temperature.

Or, if a body of  $m$  grams be made of a material whose specific heat is  $s$ , the thermal capacity of the body  $= ms$ ; and to raise it from temperature  $t_1^\circ$  to  $t_2^\circ$  requires an amount of heat

$$= ms(t_2 - t_1).$$

But, it has been remarked in the Introduction, heat is one of the forms of energy: it can be produced by the consumption of mechanical work in friction, as well as by conversion of electrical, chemical, and other energies: and under certain conditions it can, at least partially, be converted into those forms. Hence arises the problem, how much work is equivalent to a given amount of heat? or to put it concretely, how many ergs of work or other form of energy must disappear when one calorie of heat is produced? The quantity in question is known as the *dynamical equivalent of heat*, and the experiments of many physicists have been directed to showing that such an equivalent exists, and to measuring it. The solution of the problem owes more to Joule than to any one, since he first clearly stated it, and first



made measurements of an accurate character. The determination he made by the method of friction in water in 1850 was, for long, the standard experiment: besides that, he measured the heat generated by friction in mercury, friction of iron plates, compression of air, by an electric current, and in other ways, and in 1878 published a new and still more careful set of experiments by the water friction method. We shall however describe, instead of his, a research by Rowland, by a similar method, published in 1879, and generally regarded as having given the most accurate results.

Rowland's method, like Joule's, was to rotate a set of paddles vigorously in a vessel of water, hence warming the water: and to measure (i) the work spent in agitating the water, (ii) the heat generated in it, so as to be able to equate the one against the other. His apparatus is shown in Fig. 28.

The calorimeter A, or vessel in which the heat was measured, was of brass, nickel plated, and highly polished, so as to enable it to retain heat better: it was surrounded by the double-walled jacket B, which contained water between the two walls, in order that, the surroundings of the calorimeter being perfectly definite, the loss of heat from it during an experiment might be accurately estimated. Inside the calorimeter, which was of eight or nine litres capacity, was an elaborate system of vanes, attached to a spindle, working between another set of vanes fixed to the walls of the calorimeter. The spindle passed through a stuffing-box at the bottom of the calorimeter, and could be driven very rapidly by means of the gear wheels shown in the figure, from an oil-engine. The calorimeter, on the other hand, was attached by a stiff arrangement of wires to a disc, which in turn was fixed to a shaft C, passing through bearings, and suspended by a wire from the head of the apparatus. Consequently, when the engine was started the paddles on the spindle tended to drive the calorimeter round: but this being prevented from rotating, the water in it was violently churned up, and warmed in consequence. The upper part of the figure shows also the 'friction balance' by which the mechanical power spent was measured. This consisted of an accurately turned pulley D, keyed on to the shaft; round it passed a pair of tapes, leading over vertical pulleys and carrying weights. The pair of weights were so arranged as to tend to turn the shaft round in the sense opposite to that in which the paddles acted; hence by adjusting the weights correctly they could be made to balance the turning moment, or torque, of the paddles. When this balance was accurately effected there is no resultant twist on the suspending wire of the apparatus; and in order to observe whether the condition was



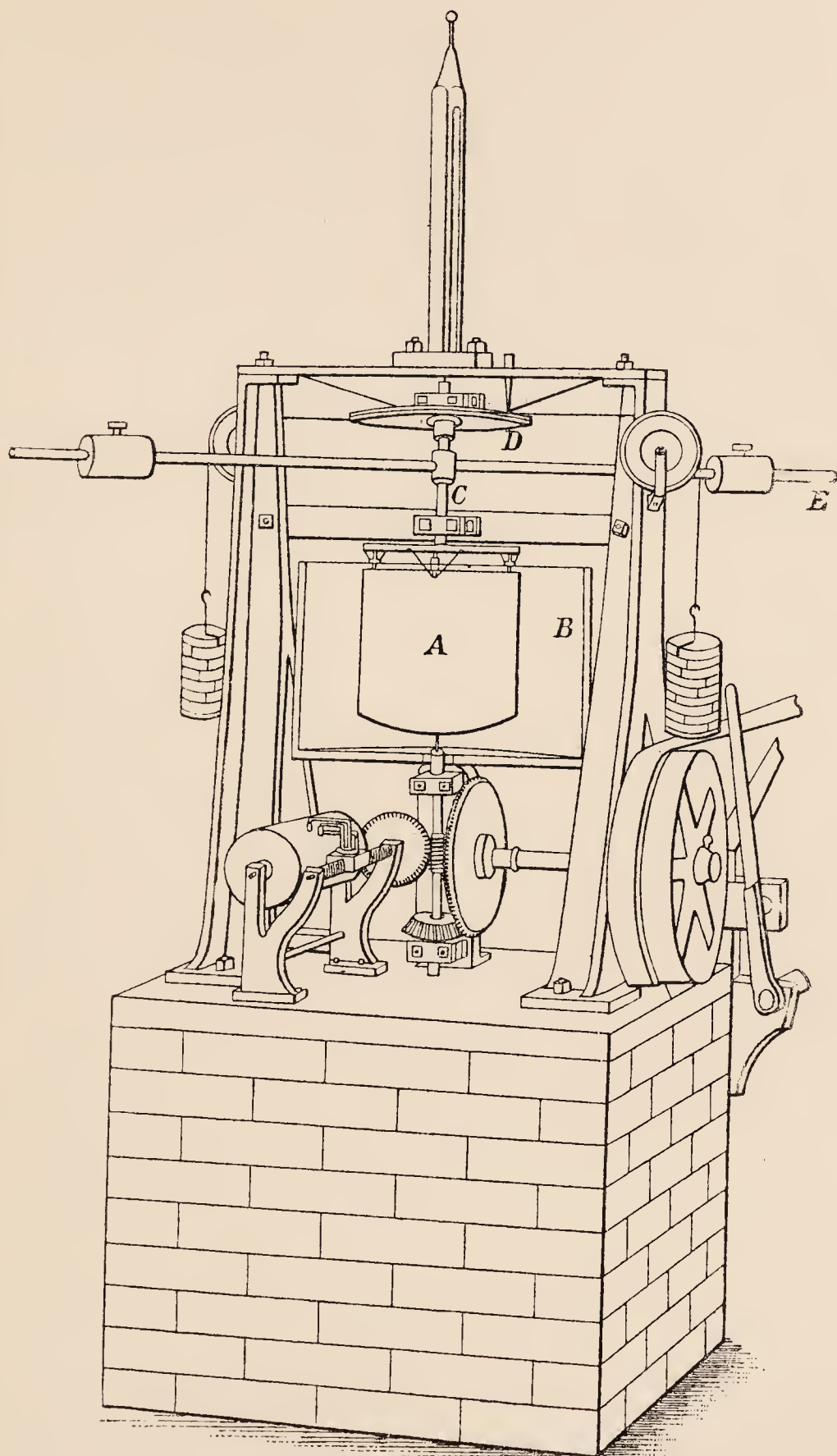


FIG. 28.

fulfilled or not, a small mirror was attached to the horizontal pulley, and a telescope focussed so as to see the reflection of a lamp in the mirror. If then the pulley be rotated even through a very small angle, the image of the lamp will be shifted in the field of view of the telescope, and by the direction of the shift the observer can tell whether the torque of the engine is overpowering the weights, or the weights are too great for the engine, and can readjust accordingly. The bar *E* with the heavy slides was for steadying the apparatus, and rendering this adjustment easier. Then the torque exerted by the weights (and therefore by the paddles) = sum of the weight  $\times$  radius of the horizontal pulley; and, as explained on p. 30, to measure the work spent by the engine we have

$$\text{work} = \text{torque} \times 2\pi \times \text{number of revolutions.}$$

The measurement of work done therefore requires, in addition to the torque, a knowledge merely of the number of revolutions: this was obtained by means of the chronograph shown in the lower part of the figure, on the drum of which the revolutions of the shaft were registered. The thermometer to show the rise in temperature of the water was enclosed in a copper case near the axis of the calorimeter, so that it could be read from outside without stopping the apparatus; and when the mercury reached each scale division, a mark was made on the chronograph, so that the number of revolutions, and consequently the work required to raise the calorimeter one degree in temperature, could be directly observed. The heat evolved was that required to raise the calorimeter and its contents in temperature. The water content was much the more important, amounting to between 8,000 and 9,000 grams, while the brass and other materials of which the vessel was made absorbed as much heat as some 350 grams more of water (i. e. the thermal capacity of the calorimeter was 350). To the amount of heat thus calculated a correction had to be made, as usually in calorimetric experiments, for heat radiated from the calorimeter; for as soon as the temperature of the calorimeter rises above that of the enclosure in which it is placed, it begins to give off heat; the amount of this heat was determined by a separate experiment. The power used in the experiments amounted to nearly half a horse-power; this is to be reckoned an advantage, as by comparison with so large an amount of heat generated, the correction for radiation is comparatively unimportant.

Rowland compared his working mercury thermometers with an air thermometer of his own construction; they have however since been standardized in a more satisfactory manner by means of the hydrogen thermometer of the International Bureau at Paris, and the values of the dynamical equivalent so corrected are probably more trustworthy than that published by Rowland himself. They are given in the following table.

Rowland's measurements extended over the range from 6° to 40°, and

showed that water (like other substances) has not quite the same specific heat at different temperatures :—

Temp.	Dynamical equivalent.	Specific heat.
6°	$4.203 \times 10^7$ ergs	1.0007 therms
10°	4.196   ,,	0.9990   ,,
15°	4.188   ,,	0.9972   ,,
20°	4.181   ,,	0.9955   ,,
25°	4.176   ,,	0.9943   ,,

It is evident then that the definition of the unit of heat previously given is not quite precise, for according to the temperature of the water it will take somewhat more or less heat to raise one gram of it through 1°. For this reason it has been proposed to adopt as the standard calorie or *therm* the quantity  $4.2 \times 10^7$  ergs. This is about the amount of energy required to heat a gram of water from 7° to 8°.

Of the numerous other methods that have been tried for measuring the dynamical equivalent of heat, only one is comparable in accuracy with the water-friction method, that is, the process in which heat is generated in a wire by the passage of an electric current through it. This method will be referred to again in the chapter on electricity, when the method of calculating the electrical energy spent has been explained.

Another experiment, however, deserves to be mentioned, not on account of its accuracy, but for its theoretical interest. Usually mechanical (or electrical) energy has been converted into heat, and the amount of heat generated measured; Hirn, however, adopted the converse plan of converting heat into work, and measuring the amounts of heat spent and work generated. There is an important distinction between the two cases however; for whereas mechanical (or electrical) energy can be wholly converted into heat, heat can only be partially converted into mechanical and other kinds of energy. Hence in an experiment of the kind now to be described it is necessary to measure three quantities of energy instead of two. Hirn's experiments were made on an actual steam engine, and therefore involved (1) the heat supplied to the engine, (2) the heat rejected by the engine to the condenser, &c., (3) the work done, which is due to the transformation of the difference between (1) and (2). In the experiments, which were carried out continuously through a working day, the pressure in the boiler was kept as constant as possible, and read every ten minutes by means of a mercury manometer; the steam from the boiler was passed through a superheater, as otherwise it would have carried with it an amount—very difficult to measure—of liquid water (priming), and the determination of the quantity of heat used would have become uncertain; as it was, the temperature of the superheated steam was taken immediately before passing into the cylinder—also every ten minutes—and its



pressure being known, Regnault's measurements of specific and latent heat sufficed to determine the heat supplied per gram of steam accurately. The mass of the steam was measured both by means of the boiler feed, and on flowing out of the condenser. The second quantity of heat, that rejected to the condenser, was determined by the quantity of condensing water and its rise in temperature; the condenser was of the 'jet' type, so that the condensed steam mixed with the water used for the condensation, and the difference in mass between the inflow and outflow of water from the condenser measured the mass of steam used. The work performed by the steam was carefully determined by means of indicator diagrams. Corrections were applied (1) for the heat radiated from the cylinder; this clearly must be subtracted from the total heat supply to get the amount actually converted into work: (2) the amount of work reconverted into heat and measured as such; this is the case with the work spent on the injection pump and air-pump of the condenser, and therefore that amount of work must be deducted to get the actual output of the engine. As an example of the results we may take the following experiment:—

Steam pressure, 4.4745 atmos.

Corresponding saturation temperature,  $145^{\circ}$ .

Actual temperature of the steam,  $19^{\circ}.57$ .

Temperature of boiler-feed,  $30^{\circ}.91$ .

Mass of steam used per stroke of piston, 198.7 gm.

This, with the known values of the latent heat and specific heat of steam (p. 106), gives as the total supply of heat

$$198.7 [606.5 + 145 \times 0.305 + (195.7 - 145) \times 0.5 - 30.91] = 128120 \text{ calories.}$$

Temperature of condensation water before use =  $16^{\circ}.15$ .

„ „ „ after use =  $30^{\circ}.91$ .

Mass of „ „ per stroke = 7732.3 gm.

Hence the heat rejected was

$$7732.3 (30.91 - 16.15) = 114130 \text{ cals.,}$$

leaving 13990 as the amount consumed; but the two corrections above-mentioned amounted between them to 1500 cal., leaving 12490 as the equivalent of the work shown by the indicator diagrams. The latter was 5318.8 kilogrammetres. Hence the dynamical equivalent is  $5318.8 \div 12490 = 0.426$  kilogrammetres, or in absolute measure

$$4.18 \times 10^7 \text{ ergs per calorie,}$$

in very good agreement with the results of Rowland's and other experiments by the direct method.

Hirn also was amongst the first to make experiments on the human body, regarded as an appliance for the conversion of heat into work. His experiments in this direction were not successful enough to quote, but others since, using more elaborate appli-

ances, have obtained results of value. For this purpose it is necessary to construct a calorimeter large enough for a man to live inside of. A recent experiment by Chauveau made in this manner gives a direct verification of the law of conservation of energy as applied to the body. A man inside the calorimeter worked a kind of treadmill, the shaft of which passed outside the calorimetric chamber and was used to drive a wheel against friction. Measurements were made on (1) the heat generated by the body inside the chamber; (2) the work done on the wheel; (3) the amount of carbon dioxide and water vapour exhaled by the man. Now if the body be regarded as a heat engine, its fuel is the tissue which, during action, is converted into carbon dioxide and water, so that from the amount of these products it is possible to calculate the heat that would be generated in producing them. In a certain instance this amounted to 257 calories per hour, while the heat actually measured in the calorimeter was 199 cals. Hence there is a loss of 58 cals. to account for the energy 'exported' from the calorimeter in the form of work. The work done on reduction by the known value of the dynamical equivalent was found equal to 68 calories, so that, considering the difficulties of the experiment, a very fair agreement was obtained; and it may be concluded that the conversion of the energy of foodstuffs into mechanical work is the same in its general character as the conversion of the energy of coal in the steam engine.

Attention may again be called to the fact that conversion of heat into work is always incomplete. The fraction

$$\frac{\text{work given out}}{\text{energy put in}}$$

for any machine is called its *efficiency*. In the case of Hirn's steam engine the fraction is

$$\frac{12490}{128120} = 9.8 \%,$$

more than nine-tenths of the total heat supply finding its way to the condenser, still in the form of heat. The best modern engines give scarcely more than 15%. The human body, according to Chauveau's experiment, is a more efficient instrument, for out of 257 cals. of energy supplied to it, it delivered 68 cals. in the form of work, or 26.5%.

### § 3. Thermodynamics.

The law of the conservation of energy, already so often mentioned, may be put in the form of a denial of the possibility of what is called 'perpetual motion.' Attempts have been made, from time immemorial, to devise a machine that should go on indefinitely turning out useful work without using up any corresponding supply of energy. If the law of conservation of energy broke down anywhere such a machine might be possible, for if a process existed in which one kind of energy was transformed into another in a ratio not equivalent, then by carrying out that process in one sense or the other energy would be gained; by an appropriately designed machine the process in question might be repeated indefinitely. That such a process has not been devised it is hardly necessary to remark: it is of more importance to point out that any new process that professed to be of such a character might safely be condemned beforehand as false, for it would be in contradiction with the whole of our knowledge of Physics: and indeed that the evidence in favour of the conservation of energy is by now so overwhelming that the impossibility of a perpetual motion may be taken as axiomatic.

Heat being a form of energy, the law of conservation is necessarily applicable to it. Nevertheless the law, as applied to heat, has often been stated separately, because taken in conjunction with another law it constitutes the basis of the mathematical theory of heat, known as Thermodynamics: it is then known as the first law of thermodynamics, and may be stated, in the words of Maxwell, as follows:—

'When work is transformed into heat, or heat into work, the quantity of work is mechanically equivalent to the quantity of heat.'

The second law of thermodynamics, which was first clearly stated by Clausius, is:—

'It is impossible for a self-acting machine, unaided by any external agency, to convey heat from one body to another at a higher temperature.'

It is, as remarked in § 1, a familiar observation that heat passes of its own accord from hot bodies to cold ones; this was indeed made the basis of the definition of temperature—temperature being at least qualitatively defined as the criterion according to which heat flows in one direction or the reverse. It might, however, be thought that by some device heat might be caused to flow indirectly from a cold body to a hot one. It can certainly be made so to pass by means of the application of work: e.g. an ice-making machine can be devised by which heat is taken from water already cooler than its surroundings and transferred to the surrounding air, the process being continued till the water freezes; but no ice-making machine will work without some external power to drive it. The statement involved in the second law is indeed not so obvious but that



numerous attempts were made to controvert it. One suggestion was that by collecting the radiation from a large hot body and concentrating it by means of powerful lenses and mirrors it might be made to raise a smaller body to a still higher temperature; this particular case, plausible as it looks, as well as others, was successfully dealt with by Clausius. The law is, of course, a generalization of such observed cases; but it is more than that, for the evidence in its favour may be immensely strengthened by a consideration of what would happen if it were not true. Just as any single exception to the law of conservation of energy would give the opportunity for a perpetual motion, and so of an indefinitely great departure from the conditions that we recognize as normal, so any exception whatever to the second law of thermodynamics would give opportunity for what has appropriately been called a 'perpetual motion of the second kind.' The atmosphere and all surrounding bodies contain vast quantities of heat in them, for we may by artificial means cool parts of them many degrees below their usual temperature, and observe how large a quantity of heat it is necessary to remove from them in order to do so. Nevertheless this energy is not in the least available for conversion into useful work: if the heat be all at the same temperature it is not possible to drive even the smallest heat-engine by means of it; in actual heat engines the heat has always to be produced at a high temperature (in a coal furnace, gas flame, &c.), in order to be capable of conversion into work. Now if it were possible by means of any exception to the second law to transfer heat from a colder to a hotter body, such means might be employed to take heat from the surrounding atmosphere and store it up at a high temperature (say in a steam boiler) in order to use subsequently for production of work; and we should have a self-acting appliance for converting the even temperature heat of the atmosphere into useful work; that is what is meant by a perpetual motion of the second kind: it involves no creation of energy, and therefore no contradiction of the first law; but it is equally contrary to experience, and shows that any exception whatever to the second law would also lead to an indefinitely great departure from observed conditions.

It is this logical consequence of the laws of thermodynamics that gives to them a validity beyond that attaching to the usual generalizations of Physics, and only comparable to that of the laws of motion. Hence any deduction made by the methods of thermodynamics has behind it all the certainty of the laws themselves. We cannot here give the mathematical expression of the laws of thermodynamics, or the deductions to be made from them; but the general character of the results may be noted as consisting in relations between the numerical values of the various properties of matter. When some property has been numerically determined by measurement it is usually possible to duplicate the result by a thermodynamic argument: e. g. if a certain chemical reaction takes place partially in the cold, then by measuring the evolution or absorption

of heat accompanying it we can foretell whether it will take place to a greater or less extent on rise of temperature.

One leading deduction from the laws of thermodynamics must, however, be noted. It refers to the absolute scale of temperature. This scale is that which would be indicated by the expansion or by the pressure of a perfect gas (p. 77) and agrees exceedingly closely with that of the hydrogen thermometer (p. 50). In the first place it may be shown that the absolute zero, on this scale, means the point at which bodies would be deprived of all their heat energy: consequently there is no lower temperature. In the next place the attainment of the absolute zero is an impossibility, notwithstanding that a somewhat near approach has already been made in the liquefaction of hydrogen and helium (at about  $-250^{\circ}$  cent. =  $27^{\circ}$  absolute). Lastly the absolute scale itself is defined by a certain relation as to the convertibility of heat into work. It may be shown that if heat be contained in a body at temperature  $T_1$  (absolute) while the surroundings are at the lower temperature  $T_2$  the fraction of it available for conversion into work is  $1 - \frac{T_2}{T_1}$ . An ideally perfect heat-engine would convert this fraction into work, leaving the remainder  $\frac{T_2}{T_1}$  in the form of heat at the same temperature as the surroundings; any actual engine will, on account of necessary imperfections, convert less.

As an example of this we may take Hirn's steam engine mentioned on p. 59. The hot body may here be regarded as the boiler, which is at a temperature  $145^{\circ}$  cent. or  $145 + 273 = 418^{\circ}$  abs.; the 'surroundings' mean, effectively, the boiler feed-water at  $30^{\circ}\cdot91$  cent. or  $303^{\circ}\cdot91$  abs. Hence the limiting efficiency of conversion is

$$1 - \frac{303\cdot91}{418} = 27\cdot3 \%,$$

whilst the actual efficiency was  $9\cdot8 \%$ .

In actual processes of all kinds inefficiency such as the above is observed; that is to say, a loss of useful work. When heat is converted into work, and anything less than the theoretical maximum of work is obtained (as is always the case in nature), an irreversible loss occurs, for the heat having become diffused at the same temperature as the surroundings is no longer available for conversion into any other kind of energy; still more is this the case when high temperature heat is simply conducted or radiated away to bodies at a low temperature, or when work is converted by friction into heat. None of these processes can be performed in the reverse sense: consequently they all mean a diminution in the amount of available energy. This fact is usually referred to by the name of the *dissipation of energy*; i. e. energy, though it cannot be lost, is continually being dissipated into a uniform state of heat out of



which it can no more be transformed. The second law of thermodynamics may, in fact, be looked upon as including two statements; first, the quantitative relations that would hold for ideal processes that were reversible in character (leading to the definition of the absolute scale of temperature); and second, that processes which are irreversible necessarily lead in the sense of the dissipation of energy.

#### § 4. Molecular theory.

Whilst the method of thermodynamics allows of co-ordinating the phenomena of heat with certainty, it does nothing towards a comprehension of the intimate nature of the transformations of energy whose quantitative relationships it allows of formulating; and the instinct of the physicist is unable to rest in so imperfect an explanation. Attempts have always been made, therefore, to trace out a description of the less obvious facts of Physics—heat, electricity, light, and so on—in terms of the ordinary, visible phenomena of mechanics. These attempts have resulted in the atomic or molecular theory of matter, and this may be regarded as another aspect of the reality, serving to complete the aspect presented by thermodynamic reasoning. The theory was suggested, in the first place, by Dalton, as an explanation of the facts of chemical combination: the law of combination in fixed proportions finds a natural explanation in the hypothesis that matter is composed of certain definite particles, *atoms*, of which there are a good many kinds—hydrogen atoms, chlorine atoms, &c.—but each kind perfectly constant in size and structure, and that these atoms unite in various ways to form *molecules*, i. e. arrangements of matter which hold together under ordinary circumstances; e. g. two hydrogen atoms unite to form one hydrogen molecule, and such molecules are the actual constituents of ordinary hydrogen gas: or a single atom may constitute the molecule, as in mercury vapour: or an atom of hydrogen unites with one of chlorine to form a molecule of hydrochloric acid—and it would be impossible to find a compound in whose molecules there was, say, only half an atom of chlorine each. The atomic theory, so started, was rendered much more definite by various physical phenomena which point to a certain measurable scale for the structure of matter. Various lines of argument, due mainly to Lord Kelvin, and to the founders of the so-called kinetic theory of gases, Clausius, Maxwell, and Boltzmann, have led to the conclusion that matter has a structure of finite, although very small size, and that just as a piece of plant tissue which looks to the eye uniform throughout, and structureless, is seen under the microscope to consist of cells of approximately equal size and similar character—perhaps a hundred of them to the millimetre in length—so all matter, water, silver, or what not, though uniform in appearance under the most



powerful microscope, is nevertheless made up of parts regularly placed and similar to one another, and that it is possible to estimate the size of these parts or molecules. The estimates of size, which have recently received valuable confirmation from the experiments of J. J. Thomson, lead to the conclusion that a molecule of hydrogen has a mass of about  $4.5 \times 10^{-23}$  grams. Since a cubic centimetre of hydrogen at ordinary temperature and pressure weighs  $9 \times 10^{-5}$  grams, this means that in 1 cubic centimetre of hydrogen there are  $9 \times 10^{-5} \div (4.5 \times 10^{-23}) = 2 \times 10^{18}$  molecules: if these were arranged in cubical order there would be  $\sqrt[3]{2 \times 10^{18}} = 1,300,000$  to the centimetre in length.

Accepting this view, then, the molecules are not to be regarded as occupying fixed positions, but as in rapid motion; and the different states of aggregation in matter are explained as follows:—in gases the molecules are some distance apart, the empty spaces between adjacent molecules being much larger than the molecules themselves. Consequently the molecules are for the most part free from one another's influence, and travel, with high velocity, in straight lines. This motion, however, brings them from time to time into collision: actually, in ordinary gases, many million times per second; the average distance that the molecules travel without a collision, the 'mean free path' as it is called, is one of the most important quantities to determine on this theory, and a knowledge of it helps to form a picture of the state supposed to exist in a gas. The molecules of oxygen and nitrogen in the air are estimated to travel at about 50,000 centimetres per second (the speed of a modern rifle bullet), and to move on an average only about

$\frac{1}{100,000}$  centimetres without coming into collision. Notwithstanding

the smallness of this distance, it is so much larger than the diameter of the molecules themselves that the interference to which their movements are subjected by the collision lasts only a small fraction of the whole time. This is still more true when the gas is rarefied, as in the vacuum tubes used for electric discharges: here the number of molecules is so far reduced that they may, on the average, travel as far as a centimetre without coming into collision. It is then possible to regard them as for the most part travelling independently, and the mathematical treatment, though difficult, is easier than in the case of liquids and solids.

In liquids, of course, the molecules are far more crowded together than in gases, since the same amount of matter goes into a space perhaps a thousand times as small. Now the phenomena of gases show that it is not sufficient to assume molecules moving freely between successive collisions, but that the molecules must be regarded as attracting one another, when not in contact, although the amount of attraction is only of importance when they are very close together. In gases this attraction only modifies a little the results calculated from the motions of

the molecules ; but in liquids, since these are always very close together, the attractions between them are always of great importance. The result is to make liquids hold together, and not, like gases, spread out into any space they have access to ; at the same time the flowing of liquids shows that their molecules must be capable of moving over one another with little difficulty. In solids the molecules are usually still closer together than in liquids, but the essential difference is in their keeping on the whole fixed positions relatively to one another ; the fact that a mark can be made on a solid, and will remain unchanged for a thousand years (as in old coins), seems to show that the molecules of the solid, though moving, only move to and fro about certain average positions, whereas those of liquids travel indefinitely through the substance.

In gases, then, the molecules, travelling with high velocity in all directions, come frequently into collision with the walls of the vessel containing the gas : and since the number of collisions per second is enormously great, the individual effects they produce cannot be observed, but the general result is a continuous pressure outwards on the wall ; this is the pressure of the gas, measured by a barometer or similar gauge, as explained on p. 37. The amount of pressure depends (1) on the number of molecules producing it, and hence is greater in a dense gas than a rare one, (2) on the velocity of the molecules to whose 'bombardment' it is due. This last point leads to the explanation of heat on the molecular theory. Heat, which is known to be a form of energy, is explained as the kinetic energy of the molecules ; it consists partly of the rectilinear movements (translation) of the molecules, just discussed, which produce the gas pressure, and partly of other kinds of movement, viz. rotations of the molecules, and, when the molecule consists of more than one atom, of relative movements amongst the atoms (intramolecular movements or vibrations of the molecule). The energy of all these movements taken together constitutes the store of heat in the gas, and in a solid or liquid the same is true, though the distribution of energy amongst the various kinds of movement may be very different. Two important theorems may, in the case of gases, be shown to follow from this idea : (1) the pressure of the gas is proportional to the average kinetic energy of the molecules of the gas ; hence when with a hydrogen thermometer we measure temperature by means of the gas-pressure we are really treating the *temperatures as proportional to the average kinetic energy of the molecules* ; (2) the average kinetic energy of translation of any one gas is equal to that of any other gas at the same temperature : consequently when there is the same number of molecules per cubic centimetre in two gases whose temperature is equal, they possess the same amount of kinetic energy of translation per cubic centimetre, and therefore the same pressure. This is Avogadro's law, stated in terms of the kinetic or molecular theory.



## § 5. Conduction of heat.

Heat, as we have seen, has a constant tendency to flow from places of high to places of low temperature till the temperature difference is abolished. This process, which is known as *conduction*, takes place through all bodies of whatever material, whether solid, liquid, or gaseous, but does not, so far as is known, take place across empty space, in the absence of matter. In order to form a quantitative conception of the process it is necessary to follow out a simple case: imagine a cubical block of a substance, 1 cm. in side, and let two opposite faces be maintained at different temperatures, one a degree higher than the other. Then at first when the temperature difference is set up, heat will flow into the substance of the block to raise its temperature; but soon each point of the block will have settled down to a steady temperature, and—the temperature of the two faces being maintained constant—a steady flow from the hot side to the cold will take place; this flow measures the process of conduction, and the number of calories per second is called the *conductivity*. The block may be looked upon as a section of a wall separating a place of high temperature from one of low, and to generalize the case we may suppose the wall to have (1) any area, then the amount of heat flowing will be proportional to the area; (2) any thickness, then the flow of heat will be inversely as the thickness; (3) any difference of temperature between its faces, then the flow will be approximately proportional to this difference. Summarizing this we have, if  $a$  is the area of the wall,  $d$  its thickness,  $t_1$  the temperature on the hot side,  $t_2$  on the cold, and  $k$  the conductivity,

$$\text{flow of heat per second} = q = \frac{ka(t_1 - t_2)}{d}.$$

The difference of temperature may here be looked upon as of the nature of a force tending to drive a current of heat: the current flowing will depend not only on the magnitude of the force but on the area, thickness, and material of the channel through which it has to flow.

To measure the conductivity of any substance, particularly of a solid, it is necessary to realize experimentally the conditions described above. Of arrangements for the purpose the earliest



was that of Péclet ; he experimented on metal plates. Each side of the plate, which was laid horizontally, was brought into contact with water—the lower hot, the upper cold, and both masses of water were vigorously stirred in order to keep the metal surfaces at the temperature of the water and avoid any stagnant layer. The conduction through the plate reduced the temperature of the lower water and raised that of the upper ; so that by observing the rate of change of temperature the amount of heat flowing through could be estimated. The measurements, however, gave values of  $k$  that we now know to be too small, and for a reason that is important to note : the two surfaces never really assumed the temperatures indicated by thermometers immersed in the masses of water, despite the most vigorous stirring. Of the whole difference of temperature available, only a part is spent in driving the current of heat through the substance of the metal : the remainder is required to cause the heat to flow from the lower water into the plate, and from the plate into the upper water. There is in fact a resistance to absorption on the one side, and to emission on the other, as well as to conduction through, and though stirring may reduce these resistances it cannot do away with them altogether : indeed, when the plate is made of such good conducting material as the metals, the resistance to conduction through the plate is so small as to be quite obscured by the others. The difficulty is not so great in the case of less good conductors, and the method has been successfully applied in the recent experiments of Lees on stones, glass, ebonite, &c. Discs of the materials, 4 cm. diameter and from 2 mm. and upwards thick, were cut, and pressed between two copper discs : one of the latter was made double, and enclosed a flat coil of silk-covered platinoid wire. An electric current being led through the wire, heat was generated in it, and flowed across the disc of material to the copper on the other side. The disc on the other side may be kept cool by a current of water. In each copper disc two radial holes are bored, inside which are soldered a platinoid and copper wire respectively ; these constitute a thermo-couple (see p. 225) by means of which the temperature of each copper disc may be measured : these are  $t_1$  and  $t_2$  in the above equation ; while the flow of heat can be determined, also electrically, from the strength of the electric current flowing through the heating coil. A little of the heat generated escapes from the edge of

the discs instead of flowing across from one face to the other, as the theory supposes, but this can be allowed for.

In the case of metals the method was successfully applied by Berget: his experiments on mercury may be taken as typical. The 'wall' or 'plate,' in order to offer sufficient resistance to the flow of heat, is thickened till it becomes a cylinder considerably longer than it is broad; this filled a vertical glass tube A (Fig. 29). The supply of heat was from a flask c in contact with the upper face of the mercury, containing alcohol, water, or other liquid, kept boiling; the current of heat flowed

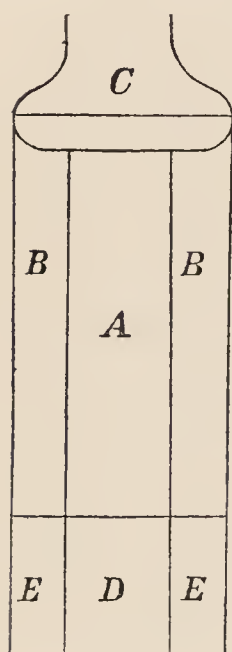


FIG. 29.

vertically downwards through the mercury to a vessel D containing ice and water in contact with its lower face. This vessel played the part of a Bunsen calorimeter (p. 105), and measured the heat transmitted. But instead of assuming, like Péclet, that the difference of temperature between the boiling liquid at the top and the ice at the bottom was all spent on the mercury, the temperature of the latter was determined by a thermo-couple that could be lowered to various measured distances in it. What is needed to know is

the quantity  $\frac{(t_1 - t_2)}{d}$ , i. e. the fall of temperature for a

given thickness. It is more convenient to regard this as a single factor in the equation—called the *temperature gradient* or *slope of temperature*; evidently this gradient, in degrees per centimetre, can be calculated if the temperature is taken at various depths in the mercury tube. If, further, we fix attention

on the flow of heat per square centimetre of cross section of the tube, we may write the equation of p. 68 in a manner that brings out more closely the essential relations:—

$$\frac{q}{a} = k \times \frac{t_1 - t_2}{d}$$

i. e. flow of heat per square centimetre per second = conductivity  $\times$  temperature gradient.

The method of experimentation just described, when the plate of material is elongated into a column or tube, increases the error due to loss of heat from the sides; this error, however, may be avoided by an ingenious device, originally used for a similar electrical problem. The device, as used by Berget, consisted in surrounding the glass tube containing the mercury by a wider concentric tube B, also containing mercury; then all the heat flowing out of the sides of the apparatus came from B, and the flow in A was left to proceed straight across from the hot to the cold face. Just as A leads below to the working chamber D of the calorimeter, so



the continuation of *B* formed the ice-jacket *E*; consequently only the heat that had flowed through *a* was measured by the calorimeter.

Of other less direct methods that have been employed to measure conductivity it is only necessary to refer to one, that depending on the propagation of waves of heat; and that not so much for the sake of the measurements made as for the light it throws on the fluctuations of temperature near the earth's surface. If one end of a bar of metal be alternately heated and cooled—say, by passing a current of steam over it for a few minutes, and then cold water for a few minutes, and so on—a wave of heat is propagated along the bar; and it may be noticed that (i) every point of the bar suffers fluctuations of temperature similar to those of the end experimented on; but (ii) it takes time for these fluctuations to travel along the bar, so that the highest temperature is reached later by other points than by the point directly heated, and later by a distant point than a near one; (iii) the fluctuations in temperature become less and less considerable as the distance from the source increases, so that if the bar is long the far end of it may not be appreciably affected. Now the earth's surface is alternately heated and cooled by the sun, passing through a cyclic change once a year, and also a cycle once a day. Accordingly the material of the earth's crust suffers similar changes; but the rate of propagation of the heat wave is so slow that only two or three metres below the surface the highest temperature is attained six months later than at the surface, i. e. midsummer occurs in January, and midwinter in July; also the extent of the fluctuations falls off very rapidly, the daily fluctuations becoming inappreciable at a depth of a metre, the annual at a few metres only. By measuring the rate of propagation and the decrease in intensity Lord Kelvin was enabled to calculate the conductivity of the soil.

The following table gives a notion of the results obtained by conductivity measurements:—

Copper	0.8	Mercury	0.018
Iron	0.16	Water	0.0013
Sandstone	0.01	Air	0.000052
Vulcanite	0.0002	Hydrogen	0.00033

This means, e. g., that a temperature gradient of  $1^{\circ}$  per cm. would cause 0.8 calories to flow per sec. through each square cm. of a copper plate. The conductivities, however, vary a little with temperature; in other words, the flow of heat is only approximately proportional to the temperature differences producing it.

The conductivity of liquids is small, that of gases much smaller; but transfer of heat can take place in them much more rapidly in another way. Rise of temperature reduces the density of fluids



in nearly all cases; consequently when a fluid is heated from below the parts first heated become lighter and float upwards, being replaced by colder fluid from above, which is in turn heated. This process is termed *convection*, because heat is conveyed by means of moving matter; it may be observed in many common instances, such as the movement of water in a kettle or flask heated over a burner, and the air over a flame, or a stand of hot water pipes. If a fluid be heated from above, convection cannot take place, as the lighter fluid is at the top to start with and stays there: this can be observed by playing on the surface of a beaker of water with a bunsen burner: the lower layers of water are heated by conduction only, and it is found that the surface water can be boiled, while that at the bottom remains cool. In measuring the conductivity of liquids it is obviously necessary to supply the heat from above in order to avoid convection. An important exception occurs in the case of water between  $0^{\circ}$  and  $4^{\circ}$ ; over this range of temperature water on heating becomes more dense, passing through a maximum of density at  $4^{\circ}$ , after which it behaves like other liquids. Hence between  $0^{\circ}$  and  $4^{\circ}$ , water, contrary to other liquids, must be heated from above and cooled from below, in order to make use of convection: an important consequence of this in nature is that lakes may be cooled in winter on the surface till the water freezes, and yet the deeper water will not fall below  $4^{\circ}$ . The same remarks are true of sea water, except that the temperature of maximum density is somewhat lower.

Convection is a process as to which it is hardly possible to make any quantitative statements, as it is so greatly affected by accidental circumstances: thus, the loss of heat from the walls of a house, or from the animal body, is much greater when a wind blows over its surface, than in still air at the same temperature; and convection in both liquids and gases can be greatly promoted by artificial circulation with pumps or fans.

## § 6. Radiation.

In the preceding section we have dealt with the processes by which heat is transferred as such from place to place; it may, however, also be transferred by means of a transformation into

another kind of energy and re-transformation, as in the process known as radiation. When heat is radiated from the sun to the earth two transformations take place; the hot matter of which the sun is composed starts waves in a medium filling the interplanetary spaces, a medium of which not much else is known than its property of transmitting such waves, and to which the name of the luminiferous ether has been given; the action, so far, is comparable with that of a ship which, in consequence of the rotation of its screw, starts waves over the surface of the sea, and provides the energy which these waves carry away with them. The waves in the ether, known generally as radiation, are propagated with great velocity in all directions. Some of them falling upon the eye produce the sensation of light, and therefore radiation of appropriate character is itself known as light. But the distinction involved in this term is a purely physiological one: other radiation, differing in no essential features from light, produces no effect on the optic nerve; but for physical purposes all radiation, whether luminous or not, must be treated together. When radiation falls on a body it is always more or less absorbed: very slightly by gases like air, very strongly by ordinary opaque solids; since then on absorption the energy of the radiation disappears, it must be by conversion into some other form, usually heat. Hence the more radiation a body absorbs, the more it will be heated; this is the second transformation above referred to. It may be compared to the action of waves on the sea, produced by the ship, falling on a small boat at a distance; they will set the boat rocking, i. e. impart kinetic energy to it, and will themselves pass on with a more or less diminished supply of energy. Conceivably an arrangement might be invented that would be so violently rocked by the incoming waves as to absorb the whole of their energy, in which case there would be no wave left to travel further, or, in other words, the absorption would be complete.

The process of communication of heat by radiation differs in its most essential features from the simple flow of heat by conduction. Conduction depends absolutely on the presence of matter to carry it on: radiation takes place best in free ether, in the absence of matter, and is always to a certain extent obstructed by the presence of matter, even in the form of gas. Conduction is a very slow process, and when occurring in waves it may take, as we have seen, months to travel a few metres:



radiation is enormously rapid, so that it only takes eight minutes to travel from the sun to the earth. These differences correspond to essential differences between the mechanism by which the two processes are carried out. According to the molecular view, conduction consists in the distribution of the kinetic energy of the molecules by their mutual collision; so that if one set of molecules have originally more average kinetic energy (i. e. are at a higher temperature) than another set with which they are in contact, molecular kinetic energy (i. e. heat) is gradually taken from the hotter and given to the colder ones by collision. Radiation, on the same view, is set up in the ether by the intramolecular movements of vibration, as we have attempted to explain by the analogy of a ship at sea; and when once such waves of radiation have been established they are propagated to a distance in the ether without further intervention of material molecules.

Intensity of radiation depends very largely on temperature, always increasing as the temperature, and therefore the store of heat-energy, increases. When two bodies at the same temperature are placed opposite one another each radiates to the other to the same extent, and consequently on the whole no interchange of heat takes place (as might be expected from the second law of thermodynamics). When one of the bodies is at a higher temperature than the other it gives out more radiation than it receives, and so on the whole loses heat.

When the difference of temperature between two bodies is small, the net radiation of heat from the hotter to the colder is proportional to the difference of temperature—a statement frequently known as Newton's law. The statement, however, may be taken to cover the entire loss of heat, or *emission* from a hot body: for the proportionality holds, as we have already seen, for conduction, and it is approximately true also for convection. Thus if a body some twenty degrees hotter than the air is exposed, it loses heat by radiation from its surface, by conduction through the air, and its solid supports, and by convection of air. All these processes take place twice as actively as if the excess of temperature were only ten degrees. The rate of emission may be measured by noting the rate at which its temperature falls: if the thermal capacity is known (see p. 55) this will evidently give the rate at which heat is being given off by the body. As the emissivity is a quantity of much practical importance, some results of measurements may be quoted with advantage:

*Emissivity.*

Black surface in vacuo . . .	0.00009
Polished silver surface in vacuo	0.00003
Black surface in air . . .	0.0002



These are the numbers of calories emitted per square centimetre per second, when the emitting body is one degree hotter than its surroundings. Generally, the flow of heat per second

$$g = h a (t_1 - t_2)$$

where  $h$  is the emissivity,  $a$  the area,  $t_1$  the temperature of the emitting surface,  $t_2$  the surrounding temperature. It will be seen from the values of  $h$  that the pure radiation (represented by the emission in vacuo) is some three times as great for a black surface as for one highly polished: and that when the surface is exposed to air, the conduction and convection add a loss of heat somewhat greater than that due to radiation from a black surface, raising the coefficient from 0.00009 to 0.0002.

When the radiating body is at a much greater temperature than its surroundings the simple law of proportionality does not hold; the amount of heat radiated increases far more rapidly than the excess of temperature. The conditions are fairly expressed by Stefan's law, that the actual radiation from each body is proportioned to the fourth power of its absolute temperature. The heat actually lost by a hot body is, as we have seen, the difference between what it radiates and what it receives by radiation from surrounding bodies. Hence we may put the loss of heat (per sq. cm. per sec.) as

$$c (T_1^4 - T_2^4),$$

where  $T_1$   $T_2$  are the temperatures of the hot body and its surroundings respectively, and  $c$  is a constant: the latter has been found for a black surface to be about  $1.28 \times 10^{-12}$  calories. The very rapid way in which the radiation increases with temperature may be illustrated by an example. If we take for  $T_2$  the atmospheric temperature, in round numbers  $300^\circ$  ( $=27^\circ$  cent.), then a body at  $327^\circ$  cent. ( $=600$  abs.) will radiate sixteen times as much heat as it receives: one at  $627^\circ$  cent. ( $=900^\circ$  abs.), a moderate red heat, eighty-one times as much as it receives, and one at  $927^\circ$  cent. ( $=1200^\circ$  abs.), or verging on white heat, two hundred and fifty-six times as much. It must be noted carefully that Stefan's law and the above calculations refer to radiation only, not to the total emission of heat, which is a complex phenomenon involving conduction and convection as well. The latter processes are more nearly proportional to the excess of temperature throughout, so that while at low temperatures they may be the most important, at high temperatures radiation, owing to its very rapid increase with temperature, completely outweighs the other causes of loss of heat.

The principles explained above throw light on practical problems of heating as involved in house construction and in clothing.

The materials of which houses are built do not vary greatly in conducting power, with the exception of thatch, walls filled in with loose shavings, and the like. These are a useful resource when heat has to be very carefully husbanded on account of intense cold outside. They are very bad conductors for the same reason that cloth, fur, wool, brown paper and similar substances are, viz. that they are largely composed of air: the air is entangled in such small spaces between the fibres of the material that convection currents are practically stopped, so that loss of heat occurs mainly by pure conduction; the heat transmitted through a layer of such material, when well made, such as fur, may be little more than that which would be calculated for an equivalent layer of air, convection being ignored. For the same reason two or more layers of a woven fabric, even when each is quite thin, keep in heat very efficiently, on account of the layers of stagnant air between them.

But as the emissivity of a surface is limited, when a current of heat is flowing through a solid wall, a part of the temperature difference available has to be spent in getting the heat away from the surface of the wall: in other words, this surface effect adds a certain amount to the resistance to conduction of the wall. Hence a wall cannot be indefinitely reduced in its resistance to flow of heat by thinning it down, or by making it of the best conducting material; it would, on the one hand, be of no appreciable advantage to make boiler plates of copper instead of steel: and on the other, a glass window, though so much thinner than a brick wall, does not let very much more heat through per square metre; and if the window be made double it may even, on account of the layer of stagnant air, serve better than the wall to keep it in.

The most efficient protection against loss of heat is a vacuum. Liquid air is now preserved in double-walled vacuum jackets, the space between the two walls being very completely exhausted of air. Here heat can only pass from the atmosphere to the liquid in the inside by radiation, and so little passes in this way as to cause only a very slow evaporation of the liquid air.

## CHAPTER III.

### PROPERTIES OF FLUIDS.

#### § 1. Perfect gases.

It is convenient to deal first with fluids (under which name both liquids and gases are included) and afterwards with solids, as the structure and properties of the latter are more complicated. In fluids, as we have seen (p. 37), there is a pressure which is everywhere and in all directions the same, except for variations due to the weight of the fluid itself, which may usually be ignored ; this pressure and the temperature are the main factors determining the properties of the fluid.

Among fluids, again, the simplest results are obtained in studying gases : with these, then, we shall begin. If a gas be kept at a constant temperature, it is found that the volume it occupies is inversely proportional to the pressure exerted on it. This is usually demonstrated by enclosing the gas in a graduated glass vessel, shut off by mercury : the mercury forms part of a gauge, by means of which the pressure on the gas may be measured. This result, known from the name of its discoverer as Boyle's law, is very closely true for ordinary gases at pressures such as that of the atmosphere ; we may express it algebraically as  $p v' = \text{constant}$  (when  $T$  is constant), where

$p$  is the pressure to which the gas is subject ;

$v'$  is the volume, which may conveniently be taken as the volume occupied by one gram, the so-called 'specific volume' ;

$T$  is the temperature.

Again, if a gas be kept in a vessel of constant volume, and heated, its pressure rises, and it is found that with all ordinary gases the rate of rise is approximately the same, viz.  $\frac{1}{273}$  of the pressure exerted at the freezing-point, per degree : this may be



shown experimentally by filling a gas thermometer, such as that shown in Fig. 25, with various gases in turn, and immersing it in ice and steam. The result obtained may, as we have seen, be most conveniently expressed by reckoning temperatures, not from the freezing-point, but from a point  $273^{\circ}$  below; using the letter  $T$  for temperatures so reckoned (and called 'absolute') we may say that

$$p \propto T \text{ (when } v' \text{ is constant).}$$

Finally, if the gas thermometer be so modified that the pressure exerted by the mercury on the gas is always the same, the gas will expand, with rise of temperature; and if the vessel it expands into be graduated we have a means of observing the relation of volume to temperature under constant pressure. It is again

found that the change for one degree is  $\frac{1}{273}$  of the amount at the

foregoing point, or, in other words, that now the volume is proportional to the absolute temperature: or

$$v' \propto T \text{ (when } T \text{ is constant) (Charles's or Gay-Lussac's law).}$$

These three results may be included in the one equation

$$p v' = R' T,$$

where  $v'$  is again to be taken as the specific volume, and  $R'$  is a constant; thus, for oxygen it is found that at  $0^{\circ}$  cent. and the normal atmospheric pressure of 76 cms. of mercury ( $= 1013230$  dynes/sq. cm.) the density is 0.0014279 gms. per c.c. or  $v =$

$$\frac{1}{0.0014279} = 700.3 \text{ c.c. per gram.}$$

$$R' = \frac{p v'}{T} = \frac{1013230 \times 700.3}{273} = 2,599,000 \text{ ergs per degree.}$$

(It should be noted that the product pressure  $\times$  volume  $= \frac{\text{force}}{\text{area}}$

$\times$  (length  $\times$  area)  $=$  force  $\times$  length  $=$  work; hence the value of  $p v'/T$  can be expressed in ergs per degree.) This equation allows of calculating the volume of a gram of oxygen at any temperature or pressure. The result may be extended to all gases in a very convenient way, by the aid of Avogadro's law; according to that well-known principle of chemistry, in gases at the same temperature and pressure, each cubic centimetre contains the same number of molecules; hence the mass of a molecule of each gas (*molecular weight*) is proportional to the mass of a c. c. (density),

or the molecular weight, taken in grams, of any gas occupies the same volume. As standard of molecular weight is taken that of oxygen as 32; the molecular weight in grams, or a gram-molecule, of oxygen occupies, at normal temperature and pressure ( $0^{\circ}$  and 760 mm.),  $32 \times 700.3 = 22410$  c. c., and a gram-molecule of any other gas occupies the same space. Hence it is more convenient to express the relation between volume, temperature, and pressure for a gram-molecule of the gas instead of a gram; calling  $v$  the volume of a gram-molecule (or *molecular volume*) we have

$$p v = R T,$$

where  $R$  is now a constant common to all gases (the so-called *gas-constant*) and its value is 32 times that calculated above for  $R'$  in the case of oxygen.

$$R = 83,157,000 \text{ ergs/degree,}$$

or very nearly 2 therms (p. 59). The equation  $p v = R T$  is called the *characteristic equation* to a gas.

The relations of pressure and volume can be most usefully represented by means of diagrams similar to that of Fig. 10, p. 23. If instead of distance and force, as in that case, volume and pressure be represented by lengths drawn horizontally and vertically, we may indicate a set of

observations relating to Boyle's law, as in Fig. 30. Thus the position of A will represent an observation, if the abscissa  $ON$  measures, on the scale of the diagram, the volume occupied by the gas experimented on, while  $OM$  measures the pressure exerted on it at the same time. Then if B, c, . . . represent in the same way other observations made on the same gas at the same temperature, we may summarize them all by drawing the smooth curve ABC through the points. ABC will therefore indicate a set of corresponding pressures and volumes, all under a constant temperature, say  $T_1$ ; the line is called the *isothermal* of  $T_1$ . If another set of observa-

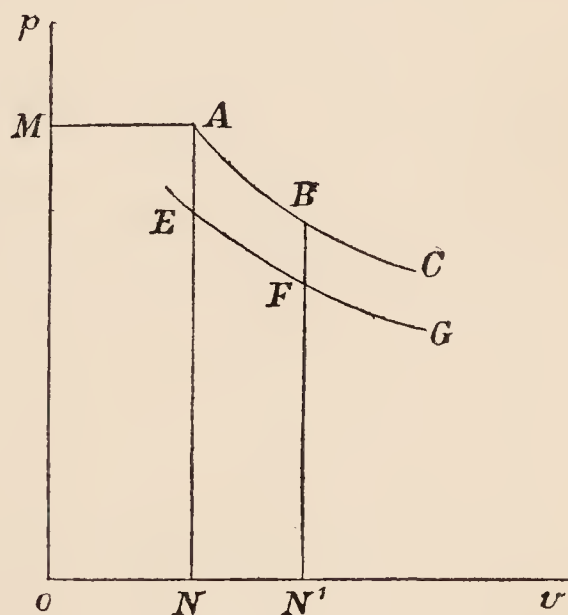


FIG. 30.

tions made on the same gas at the same temperature, we may summarize them all by drawing the smooth curve ABC through the points. ABC will therefore indicate a set of corresponding pressures and volumes, all under a constant temperature, say  $T_1$ ; the line is called the *isothermal* of  $T_1$ . If another set of observa-



tions be taken at some other temperature  $T_2$  another line will be obtained—say EFG if the latter temperature be lower than the former. This line is lower than ABC, because for any given volume the pressure exerted is less at the lower temperature : so that if A and E refer to two observations with the gas occupying the same volume, the height EN is less than AN. The curves which represent a behaviour corresponding to Boyle's law are rectangular hyperbolas, whose asymptotes are the axes of pressure and of volume,  $op$ ,  $ov$  ; but obviously the same method of graphical representation is applicable to any substance, whatever the law connecting the pressure and volume may be.

The work done by a gas in expanding may be measured with the aid of such a diagram. Since, as remarked on p. 78, the product of a pressure and a volume is a quantity of work, the  $p v$  diagram may be utilized in a manner precisely similar to the work diagram (Fig. 10), and as there shown, the work is represented by the area bounded by (i) the curve, (ii) two perpendiculars through its end points, (iii) the horizontal axis. Thus the gas in expanding from the condition of volume and pressure represented by A (Fig. 30) to that represented by B, performs an amount of work measured on the scale of the diagram by the area  $ABN'NA$ . The line joining A to B is sometimes called the path of the expanding substance, because it represents the succession of states through which the substance passes in changing from A to B. This method is by no means confined to the isothermals, whether of a gas or other substance ; if A, B are any two states, differing in temperature as well as in volume and pressure, and the line AB be drawn so as to indicate the intermediate states, the work will still be measured by the area between AB and the base and perpendiculars. Further, it may be possible to proceed from A to B by different intermediate processes, i. e. by different paths ; the work done will then in general be different, since the area of the figure will be different according to the shape of the line AB. E. g. in order to convert ice-cold water into steam at  $100^\circ$  we might either (i) heat the water to  $100^\circ$  and then boil it ; or (ii) evaporate it in vacuo at  $0^\circ$  and then heat the vapour formed up to  $100^\circ$ . The work done in the two processes would not be the same.

Another important quantity can be conveniently represented by the diagram, viz. the *elasticity*, or resistance to compression, of the fluid. A change in volume of the fluid is called a *strain* : such a change can easily be produced in a gas, but, though the fact is less obvious, liquids also can be compressed to a small extent. To produce a strain a *stress* is needed—in this case a change of pressure. Since the pressure applied to a fluid is distributed uniformly throughout it, the same pressure will clearly produce the same compression in each cubic centimetre of the



fluid, and therefore, to measure the strain, the change in volume of 1 cubic centimetre must be taken, not the change in the whole quantity of fluid, which is indefinite in amount. Now it is found that when the strains, as so defined, are small, they are proportional to the stresses required to produce them (Hooke's law), and the constant ratio of stress to strain is called the coefficient of elasticity, or

$$\frac{\text{small change of pressure}}{\text{corresponding small change of volume per c.c.}} = \frac{\text{stress}}{\text{strain}} = \text{coefficient of elasticity.}$$

Representing this graphically (Fig. 31), if  $AA'$  is a curve showing the relations between pressure and volume, we may look upon the change from  $A$  to  $A'$  as involving a stress and its corresponding strain, and if the change is small, as the figure roughly indicates, we may apply the above equation to determine the coefficient of elasticity. The stress is the increase of pressure  $MM'$ ; the strain is the decrease of volume  $NN'$  divided by the original volume  $ON$ . Hence, if  $AA'$  be produced to meet the vertical axis at  $R$ , by the properties of the similar triangles  $AA'S$  and  $ARM$  we have

$$\text{coefficient of elasticity} = MM' \div \frac{NN'}{ON} = A'S \div \frac{AS}{AM} = MR.$$

For a very small change  $AA'$  the line  $AR$  becomes the tangent to the curve; hence we may state the rule: 'To determine the elasticity of a fluid in any given condition draw the  $pv$  curve for the fluid, construct the tangent to it at the point representing the state considered, and produce it to cut the axis of pressure: the intercept  $MR$  on that axis, between the horizontal through the original point and the point of intersection, measures the coefficient of elasticity, according to the scale of pressure employed.' The pressure-volume curves discussed above were mainly under the condition of constant temperature (isothermals); if measured along them the isothermal elasticity is obtained. But the elasticity may also be measured under other conditions of compression in the same way.

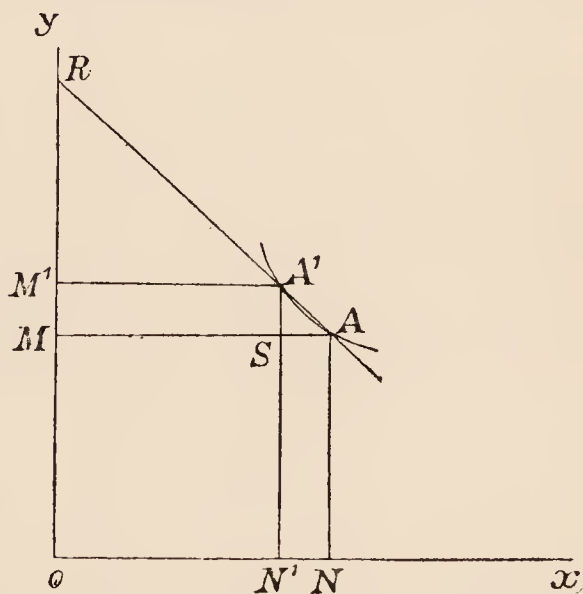


FIG. 31.

For the particular case of a gas obeying Boyle's law, for which therefore the isothermals are rectangular hyperbolas, it may be shown that the isothermal elasticity is equal to the pressure. This may be shown

algebraically as follows :—Let  $p$  be the pressure of the gas,  $v$  its volume, and let the pressure be increased by the small amount  $p'$ , causing a small decrease in volume  $v'$  while the temperature is kept constant ; then according to Boyle's law the product of pressure and volume is the same after the compression as before, or

$$pv = (p + p')(v - v') = pv + p'v - v'p - v'p'.$$

Now since  $p'$  and  $v'$  are both small quantities, their product will be much smaller still, and we may leave it out of the equation. Then we have

$$p'v = v'p.$$

But the stress (change of pressure) is  $p'$  while the strain (fractional change of volume) is  $\frac{v'}{v}$ , so that the coefficient of elasticity

$$= \frac{vp'}{v} = p,$$

the pressure of the gas. Thus the isothermal elasticity of atmospheric air is one atmosphere, or in C.G.S. units  $10^6 \frac{\text{dynes}}{\text{sq. cm.}}$ . This value may

be compared with those for liquids and solids given below. Again, if air be compressed to 20 atmospheres, its coefficient of elasticity becomes 20 atmospheres too, i. e. its resistance to further compression is twenty times as great as when at ordinary pressure.

The graphical methods above described are of general application, but the other results, viz. Boyle's and Charles' laws, the characteristic equation  $pv = RT$ , and the theorem that the isothermal elasticity is equal to the pressure, apply only to gases ; and really only constitute a limiting or ideal case to which gases approach in their properties. The term *perfect gas* is used for an imagined substance which followed these laws exactly : the nearest actual approach is made by hydrogen and helium ; somewhat less close are nitrogen, argon, oxygen, carbon monoxide, nitric oxide, methane, carbon dioxide, sulphur dioxide, ethylene, and so on.

## § 2. Condensation and vapour pressure.

When a gas such as air is compressed, it decreases continuously in volume as the pressure increases ; it may be made, by sufficient pressure, almost as dense as a liquid, but at no point is there a visible change into the liquid form. During the compression the gas does not follow Boyle's law at all closely after the pressure rises to a few atmospheres ; the actual behaviour for air may be described as follows :—At first the compression follows Boyle's law very closely, but the gas is just measurably more compressible (has a lower coefficient of elasticity) than that law indicates, so that under 10 atmospheres the discrepancy

amounts to nearly 1 per cent. If we put the product  $pv = 1$  under low pressure, then at 10 atmospheres it has fallen to about 0.992. It continues to fall, the departure from the behaviour of a perfect gas increasing in amount, till about 65 atmospheres, when  $pv = 0.97$  nearly; after this the compressibility becomes less, so that the value of  $pv$  increases again, till, under a few hundred atmospheres, the gas becomes almost as incompressible as a liquid, and the behaviour is entirely different to that of the ideal gas. Certain other gases,  $N_2$ ,  $O_2$ ,  $CO$ ,  $NO$ , and  $CH_4$ , act in a similar way, while hydrogen differs only in the fact that it is from the beginning less compressible than Boyle's law indicates, so that the product  $pv$  is always greater than unity.

It would be difficult to show these relations on a diagram of pressure and volume, since the behaviour of a perfect gas is there represented by a rectangular hyperbola, and another curve differing very slightly from that, such as would suit the observations on air, would be indistinguishable from it without accurate measurement. But if we represent the product  $pv$  vertically as a function of the pressure  $p$  (Fig. 32), then the behaviour of a perfect gas will be shown by a horizontal straight line (constant value of  $pv$ ), and that of air and hydrogen by curves, whose departure from straightness is easily noticed.

As a contrast to this, consider the behaviour of a vapour, such as steam, on compression.

To be definite, we will choose the temperature  $200^\circ$ , which is about the highest used in modern steam-engines. Steam at atmospheric pressure can exist at this temperature, if it has space enough; for if water be boiled in a flask, and the vapour be led through a tube over a flame, it can be raised ('superheated') to  $200^\circ$  or any higher temperature. If now such steam be passed into a cylinder, closed by a piston, and compressed, its behaviour is very much like that of a perfect gas; but as the pressure rises, it shows a more considerable

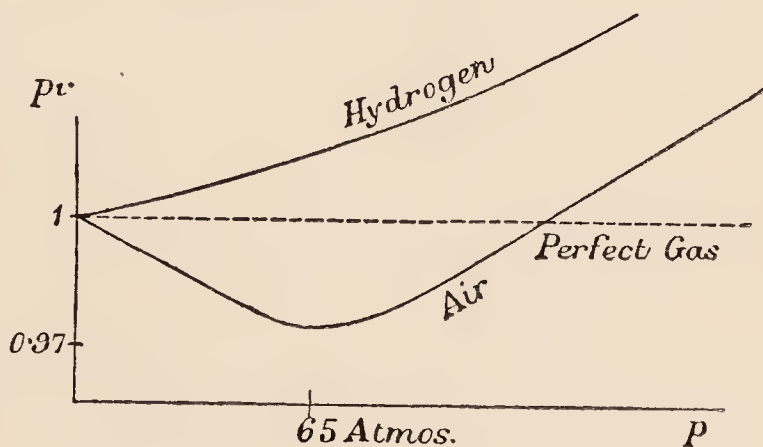


FIG. 32.



departure from Boyle's law than does air, and on the side of becoming more compressible. When, however, the piston is so far lowered that the pressure rises to 15.4 atmospheres, an abrupt change occurs: the steam begins to condense to water. Now if the piston be pressed down lower, it is found that no further rise of pressure is noted; the steam has suddenly lost its elasticity, compression being accomplished without any corresponding increase of pressure. This is observed so long as both steam and water (the two 'phases,' as they are called) are present; as the piston moves on, more and more water is formed at the expense of the steam, but the change in quantity of the phases does not affect the temperature and pressure required for equilibrium between them. Eventually, however, if the piston is continually pushed down, the volume will be so reduced that the whole of

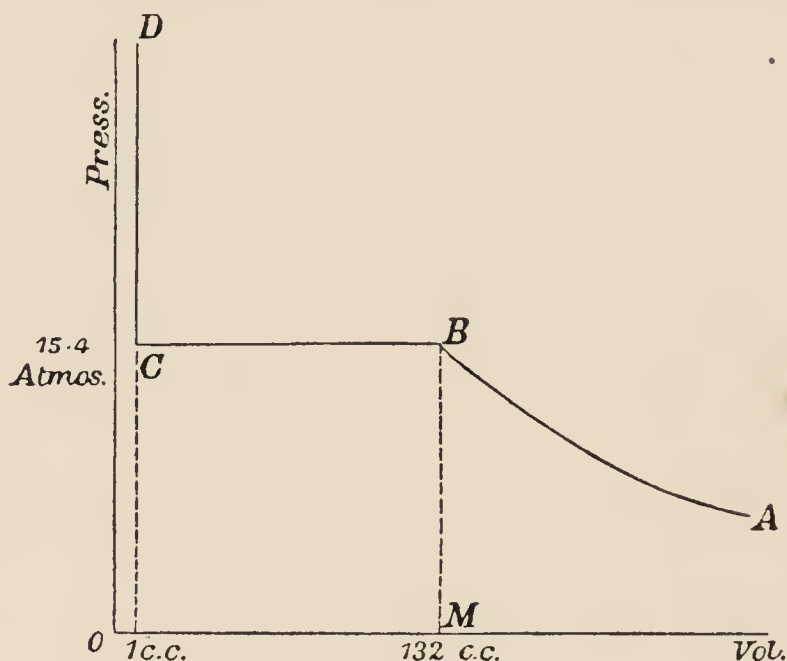


FIG. 33.

the steam is condensed; after that no reduction of volume can take place except by compression of the liquid water; then an extremely great increase of pressure is required, in proportion to the compression produced—in other words, the coefficient of elasticity is very great. These changes, which take place while the fluid is maintained at a constant temperature, characterize an isothermal for water (that of  $200^{\circ}$ ), and may consequently be shown with advantage on a pressure-volume diagram (Fig. 33, which is not to scale). The isothermal, which for air was a continuous curve, is here a broken line in three parts; the first AB,

not very different from a rectangular hyperbola, shows the compression of the vapour—rise of pressure in proportion to the diminution of volume; this part stops abruptly when the pressure has risen to 15.4 atmospheres, the volume then being 132 c.c. per gram. The steam is then said to be *saturated*, or at its maximum or saturation pressure (for the given temperature). The isothermal then proceeds horizontally along BC, indicating that the pressure remains at its saturation value until the volume is reduced to about 1.1 c.c., that occupied by one gram of liquid water at 200°; BC therefore refers to the mixture of liquid and vapour: and if the volume has any value intermediate between 1.1 and 132 c.c. (per gram), the substance will divide itself between the two phases in such proportion as to fill the space available, at 15.4 atmospheres pressure. The third part of the isothermal CD begins at the point C corresponding to complete condensation, and indicates the extremely rapid rise of pressure required to reduce the volume further; it is not a vertical line, but slopes so little towards the left that the figure does not show the slope.

According to the preceding section, the elasticity would be measured by drawing a tangent to CD and measuring the height above C at which this cuts the axis of pressure. For water the coefficient is about 20,000 atmospheres, i. e. by raising the pressure one atmosphere the volume of the water is reduced by about  $\frac{1}{20000}$  part.

The behaviour just described is that which is familiar in the case of all liquids; when heated, they are at a certain point (the boiling-point) abruptly converted into vapour—the vapour occupying a space very much greater than the liquid. The phenomenon of boiling in the air is the converse of the condensation followed out above, but as it takes place necessarily under the constant pressure of the air, it covers only the horizontal part of the isothermal. To complete the knowledge of the properties of the fluid it is necessary either (i) to trace out the isothermals for various temperatures other than 200°, or (ii) to observe the conditions of boiling under various pressures. These correspond to two different experimental methods that are in use.

(i) To trace an isothermal, with apparatus of a laboratory pattern, we may shut off the substance to be experimented on in a glass tube, by means of mercury. In Fig. 26 (p. 53) A is such a tube, graduated in order to measure the volume of the fluid; it is surrounded by the jacket B, in which a constant



temperature is maintained by the vapour of the liquid boiled in the bulb *c*, as described on p. 53. The bottom of *A* fits through an air-tight rubber stopper in the mercury vessel *E*; by means of the side tube *G* air can be pumped into this vessel to any required pressure, and the pressure measured by an attached mercury gauge. *A* is first filled with mercury and inverted in the mercury of the vessel *E*; it then constitutes a barometer. A weighed quantity of the liquid to be experimented on is then introduced into *A* in a minute stoppered flask; on rising to the top of the barometer tube the liquid drives out the stopper and evaporates, filling the space above the barometer. If then a small enough quantity of liquid has been taken, it will evaporate completely, and it will be possible to observe the volume occupied at the fixed temperature chosen, and under a pressure equal to that existing in the vessel *D*, minus that of the mercury column in the tube, i. e. it is possible to observe a point on the part of *AB* of the isothermal (Fig. 33). The pressure may then be increased by pumping more air into *D*, and another reading of pressure and corresponding volume be taken. When, however, the volume has been reduced to a certain extent the vapour will be observed to condense, and thereafter, as the substance is compressed higher up the tube *A*, no increase of pressure is required; the pressure then observed is consequently the saturation or 'vapour pressure'—that of *BC* (Fig. 33). If it be desired to continue the experimentation on the liquid, the tube *A* must be provided with a capillary ending, into which the substance may be driven as it is condensed, in order to measure the very small changes of volume that occur in the liquid state.

If a number of experiments be made on the same substance at different temperatures, it will be found that the general character of the isothermals obtained is the same as that indicated in Fig. 33, but that the saturation pressure increases greatly with rise of temperature; the following short table for water will show this:

Vapour pressure of water.				
0°	50°	100°	150°	200°
0.006	0.121	1.00	4.71	15.4 atmospheres.

In accordance with this, the least volume into which the vapour can be compressed before condensing (represented by *om*, Fig. 33) is much greater at low temperatures than at high. For this



reason the isothermal of  $200^{\circ}$  was chosen for Fig. 33; that of  $100^{\circ}$  would be still more impracticable to draw to scale, as the volume of steam at the ordinary boiling-point is about 1,650 times as great as that of water. The course of the isothermals may be further studied in Fig. 38 (vid. inf.).

(ii) The dependence of the saturation pressure on the temperature may equally well be studied by means of an apparatus for boiling a liquid under variable pressure. Such an apparatus has already been described as a means of producing constant temperatures (p. 52). If in that apparatus the enclosed barometer-tube be removed and the 'jacket' with its connexions be regarded as the essential instrument; if, further, a thermometer be placed in the jacket, we have an arrangement by which the pressure on a liquid can be varied, and the temperature of its vapour under any pressure noted. This

is sometimes known as the dynamic method of measuring vapour pressures, in contrast to that of the barometer-tube, which is known as static; the two are found to give identical results, i. e. the temperature at which a liquid boils under a given pressure  $p$  is the same as the temperature at which the saturation pressure of the vapour is  $p$ , although in the first case the pressure is due mainly to air over the liquid, while in the barometer-tube it is due to the vapour itself.

The relation between temperature and saturation pressure, which is of great importance, may advantageously be shown by a diagram in which these two quantities are explicitly represented. Fig. 34 is drawn to scale for water; it will be noted that the curve is convex to the axis of temperature, indicating that the pressure rises more and more rapidly as the temperature becomes higher; as is also shown by the table on the preceding page.

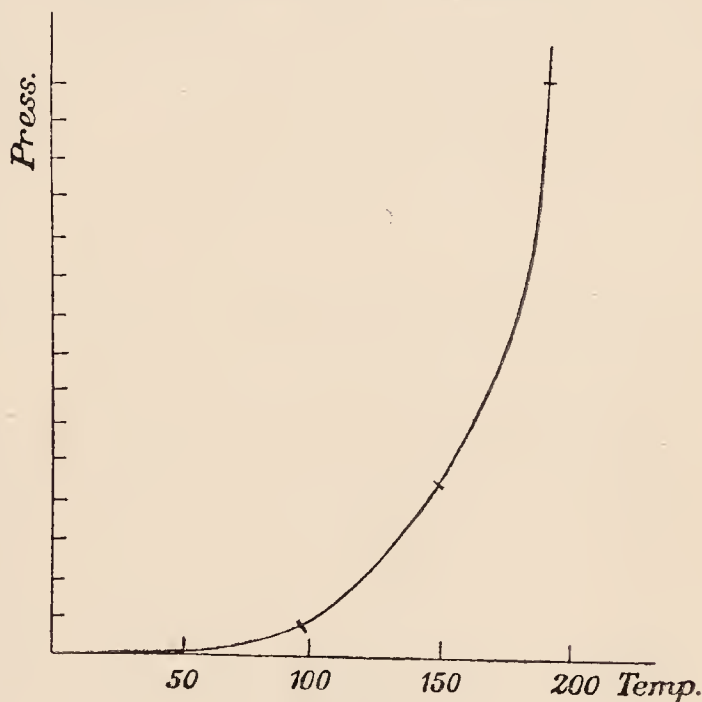


FIG. 34.

## § 3. Liquids.

Turning next to the behaviour of the substance in the liquid state, we have already noted the extremely small influence of pressure; water, which may be taken as representative in this respect, is compressed only one-twenty-thousandth part per atmosphere pressure; mercury only  $\frac{1}{250,000}$ . Temperature has a larger influence than this, although, again, less than its influence on gases; hence the most important measurement to make is on the change of volume of the liquid under the constant pressure of the atmosphere.

The change is in nearly all cases an expansion with rise of temperature, and may be expressed, as in gases, by a *coefficient of expansion*, which is defined as the increase in volume per degree, divided by the volume at  $0^\circ$ ; or, algebraically, if  $v_0$  be the volume at  $0^\circ$  and  $v_t$  that at  $t^\circ$  (cent.),

$$\text{coefficient of expansion} = \frac{v_t - v_0}{t v_0};$$

but whereas in gases this coefficient is always about  $\frac{1}{273}$  it varies much from one liquid to another.

The coefficient of expansion may be measured by any of the processes available for measuring the density of liquids (p. 10), if carried out at various temperatures, since it merely expresses the change in density produced by temperature. It may, however, best be accomplished independently by a *dilatometer*, which is constructed after the principle of the thermometer. The instrument is shown in Fig. 35; it is made entirely of glass, consisting of a wide tube A, to which are sealed the capillaries B and C; B has a number of small bulbs blown on it, and marks etched between each, C is bent as shown, and drawn out very fine at the end. After being cleaned and dried it is filled with mercury, by attaching the end of B to a filter-pump and dipping C under mercury in a dish; the mercury is allowed to rise to each of the marks in turn, and the

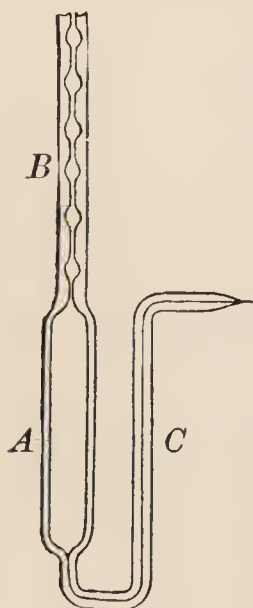


FIG. 35.

weight taken: from this the volume of the main tube and each of the small bulbs can be found. The apparatus is next filled with the liquid to be experimented on (previously freed from dissolved air by boiling) to a convenient level; the point of c is fused off in the blowpipe, B closed by a cap of rubber tubing, and the whole placed in a bath of water alongside a thermometer. The bath is heated and kept well stirred, and note taken of the temperature at which the liquid reaches each of the marks. The experiment may with advantage be extended to temperatures below that of the air, by means of freezing mixtures, the dilatometer and thermometer being placed in a bath of toluene or paraffin oil. The results may be plotted to form a curve, with temperature and volume represented along the two axes; from this the volume at the freezing-point  $v_0$  may be found, and the coefficient calculated according to the formula on the preceding page.

It will usually be found that the increase in volume is not quite the same for each degree rise of temperature; when that is the case

the entire result cannot be expressed by means of a single 'coefficient,' for the coefficient will vary according to the temperature chosen for calculation. We must then distinguish between (i) the *mean coefficient of expansion* over a given range of temperature  $t_1$  to  $t_2$ , or

$$\frac{\text{volume at } t_2 - \text{volume at } t_1}{(t_2 - t_1) \times \text{volume at } 0^\circ}$$

and the *true coefficient of expansion* at a given temperature  $t$ , which is the value of the coefficient derived from observations very slightly above and below  $t$ . This distinction is most clearly shown graphically, as in Fig. 36. If the expansion of a liquid were the same for each degree, i. e. the coefficient of expansion constant, it would be represented by a straight line, and the steepness of slope would

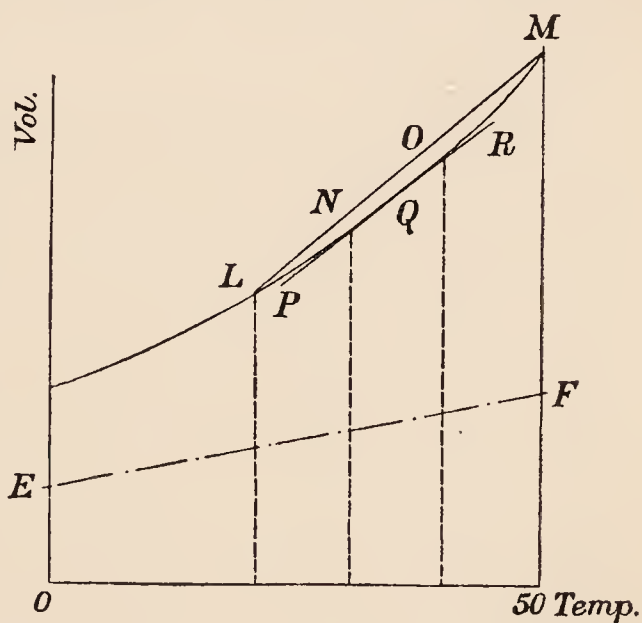


FIG. 36.



be a measure of the coefficient; if, for instance, the volume at  $0^\circ$  and  $50^\circ$  be known, the difference between them (vertical distance between E and F) must be divided by the difference of temperature (horizontal distance between E and F) to give the rise per  $1^\circ$ , so that the quotient, i. e. the tangent of the angle of slope, is proportional to the coefficient. Next, in the actual case, to get the mean coefficient between  $20^\circ$  and  $50^\circ$  we must join L and M, and the tangent of the slope of this line will be proportional to the quantity required; or if the mean coefficient between  $30^\circ$  and  $40^\circ$  be desired, it will be given by the slope of the chord NO, while in the limit the chord becomes a tangent, and the true coefficient of expansion at  $35^\circ$  is indicated by the slope of the tangent at that point, PQR.

The coefficient of expansion obtained by means of a dilatometer (whether true or mean coefficient) requires a correction on account of the expansion of the glass vessel itself; since this expands with rise of temperature, the contained liquid will not rise so far up the stem as if it did not; hence the coefficient of expansion, as observed, will be too low. The correction amounts to about 0.000025 for ordinary glass, so that e. g. the coefficient of expansion of benzene as observed is 0.001176 per  $1^\circ$ ; the true value 0.001201. The coefficients for different liquids vary pretty regularly according to their boiling-points, being smaller the higher the boiling-point; the value just given for benzene may be taken as an average for ordinary liquids; mercury, in accordance with its much higher boiling-point, has a coefficient as low as 0.000180 per degree.

The behaviour of water is peculiar. From  $0^\circ$  to  $4^\circ$  it contracts with rise of temperature; its volume is then a minimum, and at  $8^\circ$  is about the same as at  $4^\circ$ . The expansion becomes more rapid as the temperature rises, but even at  $100^\circ$  is abnormally small; the curve representing its behaviour consequently slopes much less steeply than that of a normal liquid, such as benzene. The physiological effects of this peculiarity were referred to in the preceding chapter.

Of methods other than the dilatometer for measuring the expansion of liquids, we may mention the following:—

(i) The weight thermometer: this is essentially similar to the specific gravity bottle in principle; it consists of a bulb—say 10 c.c. in capacity—drawn out to a fine point. It is weighed empty, filled by alternate

heating and cooling, in much the same way as a thermometer, with the liquid to be studied, and when completely full at a known temperature (that of a water bath or ice bath) weighed again; it is then heated to some higher temperature, when the expansion drives out a little of the liquid, and on cooling weighed again. Regarding the volume of the bulb as unchanged in this process, it is clear that the weight of liquid contained will be proportional to the density, therefore inversely as the specific volume of the liquid. If the latter be  $v_0$  at the freezing-point, and the coefficient of expansion  $\alpha$ ,

$$\frac{\text{Weight of liquid contained at } t_1}{\text{Weight of liquid contained at } t_2} = \frac{\text{sp. vol. at } t_2}{\text{sp. vol. at } t_1} = \frac{v_0 (1 + \alpha t_2)}{v_0 (1 + \alpha t_1)} = \frac{1 + \alpha t_2}{1 + \alpha t_1}$$

whence  $\alpha$  can be calculated. Here again the coefficient  $\alpha$  obtained is that relative to glass, and must therefore be increased by 0.000025 to give the true value.

(ii) The areometric method: this consists in weighing a solid in air, and when immersed in the liquid as described on p. 12, and repeating the process at different temperatures. The apparent loss of weight of the solid is proportional to the density of the liquid in which it is immersed; hence this method, like the last, is essentially a method of comparing the density at different temperatures, and of course gives results relative to the solid used, and must be corrected by adding the coefficient of expansion in volume of the solid, as before.

(iii) Just as the densities of two liquids may be compared by arranging a column of each to produce the same hydrostatic pressure, so the density of the same liquid at different temperatures may be studied. Regnault in this way measured the coefficient of expansion; his apparatus consisted of a pair of vertical tubes AA, BB (Fig. 37) to contain the mercury, one immersed in a bath of tap water, the other in a liquid bath that could be heated: the two tubes were connected at the top by a cross-tube of fine bore AB, which served to equalize the pressure on the two sides, but allowed so little liquid to diffuse through it as to cause no appreciable transfer of heat from the hot to the cold side. At the bottom each of the main tubes was connected to a short upright glass tube, c, d; these were connected at the top to one another and to a supply of compressed air. c and d served as a gauge to measure the difference in pressure produced by the two columns A and B of mercury, which were of the same length but different densities; since (p. 36) the pressure produced by a liquid column is proportional to its density, this gives the means of determining the effect of temperature on the density

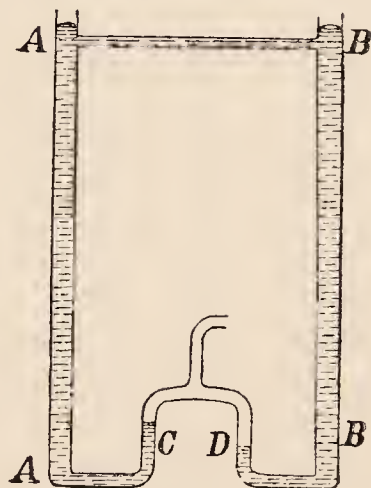


FIG. 37.

(and therefore volume) of mercury ; and since the pressure produced at a given depth is independent of the size or nature of the containing vessel, this method, unlike the preceding ones, requires no correction for the expansion of glass, but gives directly and accurately the coefficient of expansion of the liquid studied.

#### § 4. Critical point.

The marked difference, noted in § 2, between the behaviour of air and steam on being compressed, is not characteristic of those substances, but according to the temperature either set of phenomena can be obtained ; steam at a very high temperature (over

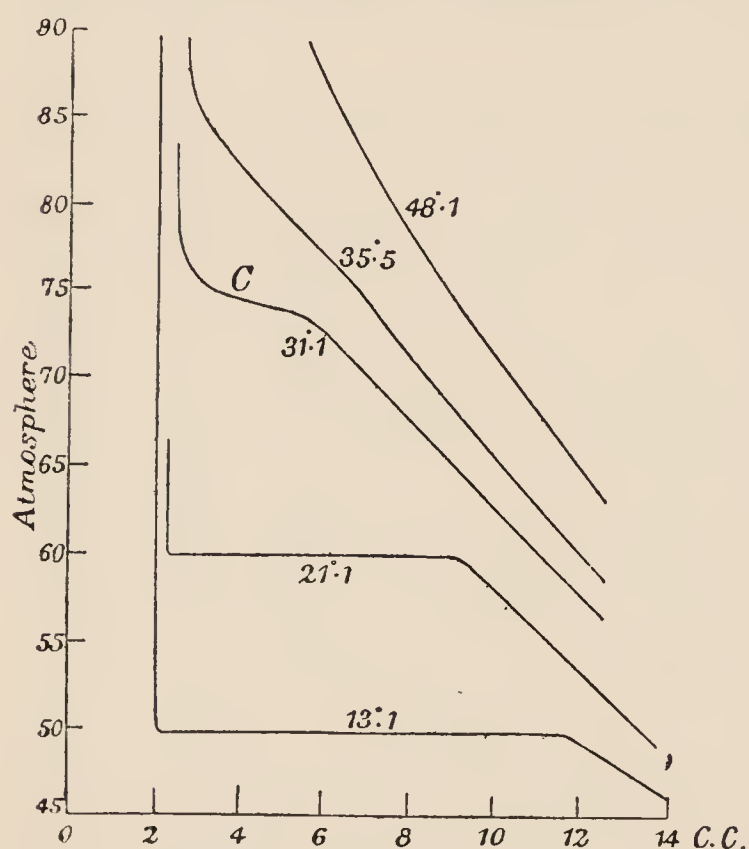


FIG. 38.

400°) would behave on compression like air under ordinary conditions, while air at  $-160^{\circ}$  would show the phenomena described above in the case of steam. Neither is a convenient material on which to show the transition between the two states on account of the extreme temperatures required : carbon dioxide is more suitable, and was indeed the substance with which Andrews made such experiments and discovered

the existence of the critical point. It has already been remarked in the case of steam, that with rise of temperature the saturation pressure rises greatly, and so the volume occupied by a gram of saturated vapour becomes less and less (Fig. 38). This is well brought out by the behaviour of carbon dioxide : the gas was compressed into a strong capillary glass tube, and in an experiment made at  $13^{\circ}.1$  found to condense under about 50 atmos. pressure ; this then is the saturation



pressure of the substance at that temperature, the behaviour being similar to that previously described for steam. But on account of the high pressure involved the specific volume of the saturated vapour was only 12 c.c. whilst that of the liquid formed was 2 c.c. But at  $21^{\circ}.1$ , the next temperature used in the experiments, the approximation between vapour and liquid had gone still further: the vapour pressure was 60 atmos., the volume of the vapour before condensation 9 c.c., that of the liquid formed about 2.2. Finally at  $31^{\circ}$  the distinction into vapour and liquid ceased: at no time during the compression were two forms of the substance visible in the experimental tube together, as was the case at lower temperatures, and at higher temperatures the behaviour was analogous to that described for air. The temperature limit between these two conditions is called the critical temperature—it may be looked upon as the limit beyond which the phenomena of evaporation and condensation cease to exist, or beyond which liquid and vapour cease to be distinguishable. The saturation, or vapour pressure, therefore, is a quantity relating only to temperatures below the critical, and the limiting value corresponding to that temperature is called the critical pressure. The state of the fluid marked by the critical temperature and pressure is called the *critical point* (c, Fig. 38), and the volume the fluid then occupies, the critical volume; the latter is, clearly, at the same time the least volume of the saturated vapour and the greatest volume of the liquid.

There are only a few substances whose critical points lie below the atmospheric temperature, and consequently cannot be condensed by pressure alone: they include hydrogen, oxygen, nitrogen (and consequently air), carbon monoxide, methane, nitric oxide, argon and helium. Two or three distinct methods have been successfully employed to cool these gases sufficiently to permit of their being condensed. That of Pictet was to proceed by steps: sulphur dioxide was liquefied by compression, and the liquid rapidly evaporated by means of a pump, the liquid thus fell in temperature to its boiling-point under the low pressure maintained by the pump, viz. about  $-70^{\circ}$ ; the liquid and vapour rising from it were used as a jacket to maintain this low temperature, on the same principle as the jackets described, p. 76; by means of it carbon dioxide, under a pressure of a few atmospheres, was condensed and allowed to collect in a separate

reservoir, from which it was rapidly evaporated by a second pump. The boiling-point of the carbon dioxide was thus reduced to  $-130^{\circ}$ : it was used as a jacket for the tube in which oxygen was to be condensed, and on forcing oxygen at high pressure into this the liquefaction was accomplished, showing that the critical point of oxygen is not so low as  $-130^{\circ}$ . Cailletet liquefied oxygen at about the same time as Pictet, with simpler apparatus, by adopting a novel principle; the gas was raised to as high a pressure as possible in an apparatus of the same essential character as that referred to on p. 86, but very strongly constructed, so as to withstand pressures of 1,000 atmospheres: the glass tube containing the gas was enclosed in a steel block, from which only the capillary end projected; the gas when compressed was cooled as far as could conveniently be done by a bath of liquid sulphur dioxide, still it did not condense. Then, by opening a valve, the pressure was suddenly reduced to a few atmospheres only: the gas in expanding does work in driving the mercury or water compressing it out through the valve; this work is done at the expense of the store of heat-energy in the gas itself, which consequently falls in temperature. The fall in temperature can be made great enough to condense oxygen or nitrogen, which appear in small drops on the inside of the capillary tube: but the quantity obtained is very small and does not allow of any practical use being made of the liquid. A further modification of principle, introduced quite recently by Linde, allows of the liquefaction being carried out continuously, so that any quantity can be obtained: liquid air can by that means be prepared in a few minutes, and at a cost of one or two pence per litre. When air under pressure escapes through a fine hole, e. g. through a throttle-valve, the work it does in expanding is spent in overcoming friction in the valve, and consequently the heat is generated there and returned to the expanding gas: its temperature therefore does not fall like that of a gas expanding in such a way as to do work on an external system, as in Cailletet's experiment, or by means of a cylinder and piston; indeed if the gas were a 'perfect' one it would issue from the valve at precisely the temperature it went in at. In air, however, the two effects do not quite balance: the heat lost by the gas on account of the expansion being somewhat greater than that returned to it by friction (as the equivalent of the work done by the expanding



gas); hence there is a small fall of temperature, about a quarter of a degree per atmosphere pressure. In machines constructed after this principle a pressure of some 200 atmos. is used: this would not in itself cool the air sufficiently to condense it, but the cooled air coming from the valve is passed through a pipe called an 'interchanger,' which jackets the air approaching the valve and cools that, so that the process intensifies itself, till eventually a part of the air liquefies. The method consists, therefore, in compressing air by a machine pump to about 200 atmos., passing the compressed air through a spiral tube cooled by tap-water, to remove the heat generated in the compression, freeing it from carbon dioxide and water vapour by means of soda-lime, and letting it flow through the interchanger to the throttle-valve and back; when the apparatus has been working for a few minutes a portion of the air liquefies at the valve and drops into a receiver. The latter is a double-walled glass vessel with a vacuous space between the two walls (cf. p. 76). Air liquefies under the atmospheric pressure at  $-194^{\circ}$ . Hydrogen can be liquefied in the same way if previously cooled by liquid air: its liquefying-point (or boiling-point) is about  $-235^{\circ}$ .

### § 5. Mixtures.

When two or more gases are mixed, they each produce almost the same effect as if they separately occupied the entire space. In particular they each produce a partial pressure equal to that which would be produced if the other gases were not present (Dalton's law); the total pressure of the gas, as measured by a manometer, being the sum of these partial pressures. Thus air is a mixture of oxygen (21 % by volume) and nitrogen (79 %); if the two pure gases were taken in these proportions, and made to occupy in turn a glass vessel of fixed volume, and their pressures measured by an attached gauge, and then the two gases mixed and compressed into the same vessel, the pressure would then be found to be equal to the other two measured pressures added together. Further, a vapour, whether saturated or unsaturated, in presence of gases, produces its own partial pressure, just as if the gases were not there. We have already had an instance of this in the agreement between the static and dynamic methods for measuring vapour pressures (p. 85), since the former measures



the vapour pressure of a liquid in the absence of any other substance, while the latter does so in presence of air.

This principle is of importance in finding the mass of a specimen of gas. Gas is usually collected for measuring in graduated glass vessels over mercury, water, dilute sulphuric acid or other liquid. Mercury has no appreciable vapour pressure at ordinary temperatures, but water and dilute acid have; hence the pressure as measured is that of the gas saturated with water vapour, and to get the pressure of the gas alone, that of the vapour must be deducted. Thus, suppose  $b$  to be the height of the barometer, and that the liquid over which the gas is collected stands at a height  $h$  higher in the collecting vessel than in the trough outside, while  $d$  is the density of the liquid: then the column of liquid inside the collecting vessel produces a pressure equal to

that of a mercury column  $\frac{hd}{13.6}$  high (13.6=density of mercury);

so that the actual pressure inside the collecting vessel is  $b - \frac{hd}{13.6}$ .

If at the temperature of the experiment ( $t^\circ$  cent.) the liquid has a vapour pressure= $s$ , the partial pressure of the gas is  $b - \frac{hd}{13.6} - s$

( $s$  may be taken from published tables; it is less for dilute acid than for pure water). Now the density of a gas is by Boyle's law proportional to its pressure, and by Charles's law inversely proportional to the absolute temperature  $T$  ( $=t+273$ ); hence if  $\rho$  is the density of the gas at  $0^\circ$  and 76 cm. pressure, its density under the observed condition

$$= \rho \times \frac{b - \left(\frac{hd}{13.6}\right) - s}{76} \times \frac{273}{273+t},$$

and the mass of gas is obtained by multiplying this quantity by the observed volume. When the gas is collected over mercury or other non-volatile liquid, such as strong sulphuric acid, the term  $s$  can be neglected.

The air expired from the lungs is also saturated with water vapour at the temperature of the body or nearly so; the vapour pressure of water at that temperature ( $37^\circ$ ) is 47 mm. Hence the water vapour in the breath amounts to  $\frac{47}{760}$  of the whole (by volume). Carbon dioxide is found to constitute about 2.2 % or 17

parts in 760, and consequently exerts a partial pressure of 17 mm. It may be desired, by ventilation, to reduce the partial pressure of the carbon dioxide to 1 mm. (in fresh air it amounts to 0.3 mm., so that this may be regarded as reasonable); to do so it would be necessary to make the volume of the air 17 times as great, or rather more so, seeing that the ventilating air itself contains a little. If the volume of air expired be taken at 8 litres per minute per head, it may be reckoned that about 150 litres a minute is a sufficient supply.

Atmospheric air always contains some moisture, and for meteorological purposes it is desirable to estimate the amount of it. The most direct process is to draw a measured volume of air through drying-tubes and weigh the moisture absorbed; this, sometimes known as the *chemical method*, is represented in Fig. 38a.

The aspirator A serves to draw a measured volume of air through the tube D into the drying bulbs C: these may contain calcium chloride, strong sulphuric acid, or phosphorus pentoxide, of which the last is the most efficient; the drying bulb B serves to keep

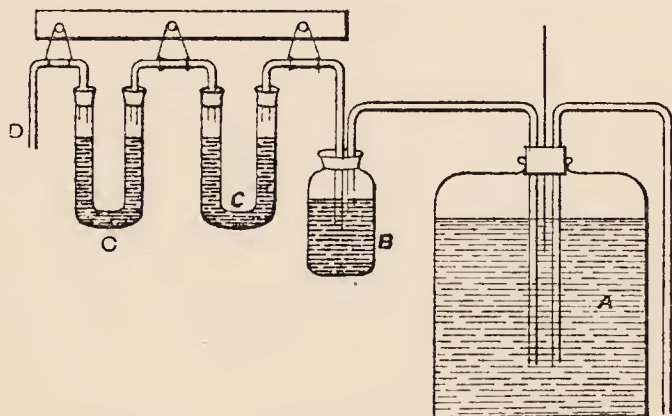


FIG. 38 a.

any vapour from the water in the aspirator from diffusing back into the bulbs C. From the increase in weight of C the concentration of the vapour may be calculated, and thence, if the temperature of the air be noted, the pressure obtained. The results may be expressed either as (i) actual partial pressure (incorrectly called 'tension') of the water vapour; (ii) the ratio which that pressure bears to the saturation pressures at the same temperature (called the relative humidity); (iii) the concentration of the vapour (weight in unit volume); (iv) the drying power, i. e. the additional weight that unit volume of the air could take up before becoming saturated. All these are recorded in the daily weather reports as published in the *Times*.

A more rapid process to determine the hygrometric state of the air consists in determining the *dew point*. Dew is deposited on any surface that is cooled till the air in its immediate neighbour-



hood becomes saturated with vapour. To determine the dew point a polished metallic surface (usually a small silver cup) is cooled, by evaporating ether inside it, and watched till dew forms on it ; it is then left to itself till the dew disappears by evaporation, and the temperature at that moment noted ; we may then find from the tables of vapour pressure of water what the pressure corresponding to the observed dew-point temperature is ; that will be the actual pressure of the water vapour in the air.

Another process is that of the wet and dry bulbs, i. e. of reading the temperature of the air by an ordinary (dry) thermometer, and reading at the same time another thermometer whose bulb is kept wet by a wrapping of cotton wool dipping into a vessel of water. The process is not reliable, however. A more useful instrument is the hair hygrometer : this consists of a long hair, previously freed from grease by means of ether, which is kept stretched by a small weight ; it contracts in moist air and expands in dry, and may be made to record its changes in length by means of a magnifying lever working over a scale. The instrument is only roughly accurate, but is very useful for certain observations, because it is direct reading. Its scale must be calibrated by comparison with determinations of humidity made by the other methods.

When two liquids both give off vapour to the same space, it is still true that the pressure produced by each vapour is that that would be calculated from the quantity present (the concentration) independently of the other. But it is not in general true that the saturation pressure of either remains unaffected ; if the two liquids do not mix at all, then each has the same vapour pressure as if the other were not present, and the total vapour pressure is the sum of the two ; this is nearly the case for benzene and water. If the two liquids are completely miscible, the vapour pressure of any mixture usually lies between the pressures of the two pure substances, approximating to that of either component according as that component preponderates in the liquid ; this is true, e.g., of benzene and carbon tetrachloride. Sometimes, however, a mixture of two liquids in certain proportions may possess greater vapour pressure than either separately (benzene and alcohol), or less than either separately (water and formic acid) ; the latter case indicates a considerable attraction between the molecules of the two substances, and the mixture may be regarded as approaching a chemical combination in its properties.

When a solid is dissolved in a liquid, the solid being practically non-volatile, the only vapour pressure to be considered is that of the solvent,



and that is diminished. The lowering of vapour pressure produced can be calculated according to a rule first given empirically by Raoult, and subsequently explained on theoretical grounds by van't Hoff. If the solution contains  $n$  molecules of dissolved substance for one of solvent, the vapour pressure is  $1 - n$  times that of the pure solvent. This law, however, like other generalizations with regard to solutions (vid. inf. p. 127), applies only when the solution is sufficiently dilute, just as Boyle's law is applicable to gases only when their concentration (or density) is small enough. As an example we may take a 1 % aqueous solution of cane sugar  $C_{12}H_{22}O_{11}$ ; the molecular weight of sugar is  $(12 \times 12 + 22 \times 1 + 11 \times 16 =) 342$ ; that of water  $H_2O$  is 18. Hence 1 gram molecule (18 grams) of water contain  $\frac{18}{100} \times \frac{1}{342} = \frac{1}{1900}$  gram molecule of sugar; therefore at  $100^\circ$ , while the vapour pressure of pure water is 760 mm., that of the sugar solution is less by  $\frac{1}{1900}$  part, and amounts to 759.6 mm.

Raoult's rule has been used to determine the molecular weight of bodies in solution; it leads in some cases to apparently anomalous results, the most important being for solutions of salts, acids, and bases in water. Here the vapour pressure is lowered more than would be expected according to the rule, and these observations, with other evidence, have led to the 'electrolytic dissociation theory,' according to which salts in solution are dissociated; the number of dissolved particles is therefore greater than would be the case if the salt were not dissociated, and the change of vapour pressure correspondingly increased.

To say that, at a given temperature, the vapour pressure of a solution is less than that of the pure solvent is equivalent to saying that, under a given pressure, the boiling-point of a solution is higher than that of the pure solvent, since in all cases the vapour pressure increases with the temperature, and the boiling-point is the temperature at which the vapour pressure becomes equal to the external pressure on the liquid. It is convenient in practice to measure the rise in boiling-point, rather than the diminution in vapour pressure; the measurement is usually made in an apparatus such as that shown in Fig. 39, the design of which is due to Beckmann. It consists of a boiling tube A provided with two side tubes  $t_1$ ,  $t_2$ ;  $t_1$  serves for the introduction of the substance,  $t_2$  carries a condenser K. The boiling tube stands on a sheet of asbestos L and a piece of wire gauze D; it is also supported by a clamp N. As an air jacket to the boiling tube, the open cylinder of glass G is placed round it, and surmounted by the sheet of mica S. C is a calcium chloride tube, to be used with hygroscopic materials, or when the cooling water in the condenser would cause a deposition of dew. The condenser may be dispensed with in the case of liquids boiling above  $100^\circ$ . The thermometer is one of high sensitiveness and adjustable range, as described on p. 49. The solvent may be introduced by a pipette or weighed out in the boiling tube, which can be hung to the arm of a balance; there must be enough of the liquid to cover the bulb of the

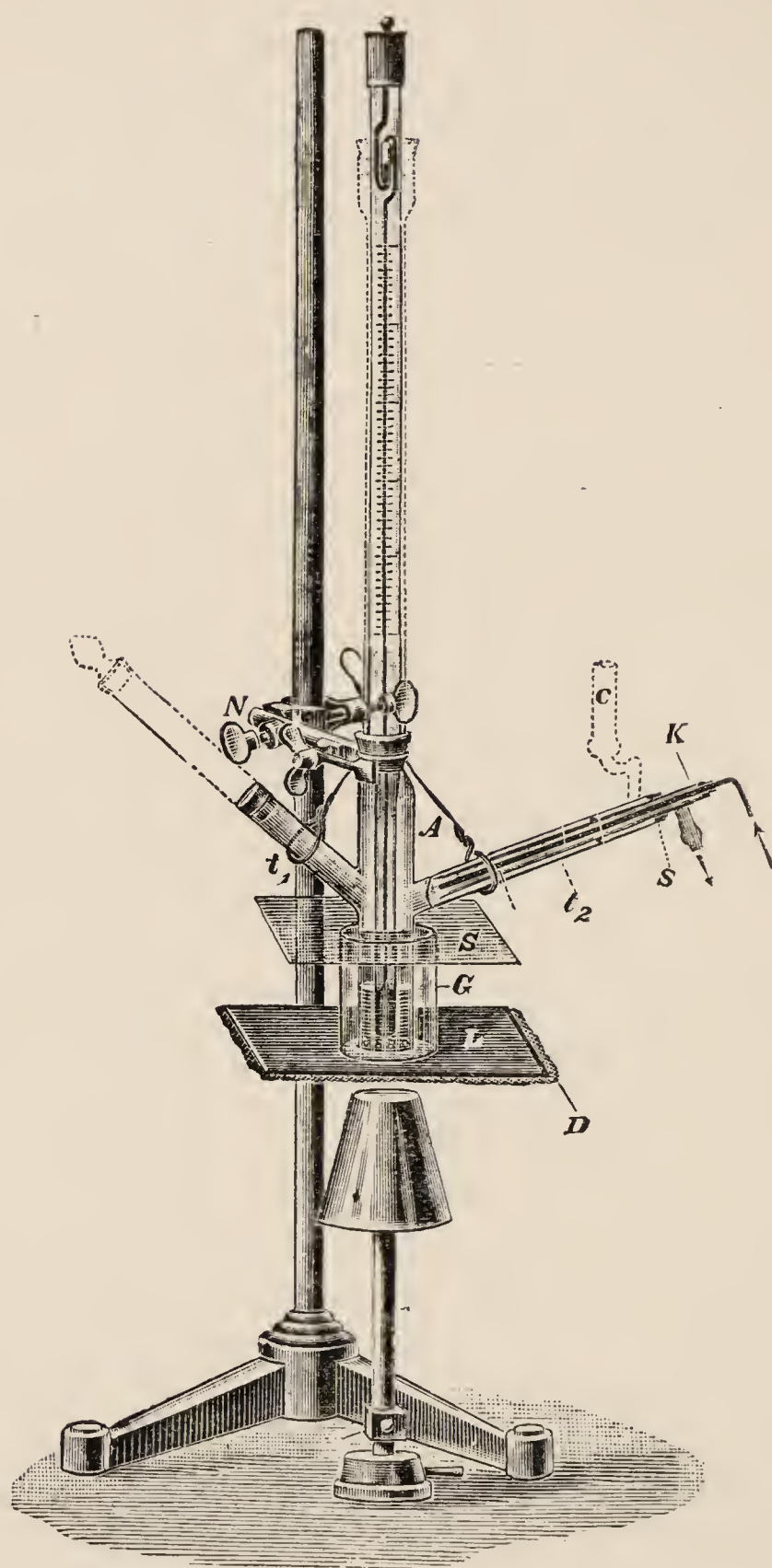


FIG. 39.

thermometer, for which purpose 10 c.c. suffices. The substance to be dissolved may conveniently be introduced, if solid, in the form of a small pastille, shaped in a steel press; if liquid, by means of a small pipette somewhat similar in shape to a Sprengel pyknometer. The liquid will not boil freely and impart a constant temperature to the thermometer unless some such material as glass beads or platinum foil be placed in the boiling tube; these supply minute air bubbles which start the formation of vapour bubbles in the liquid: without them the liquid is apt to rise in temperature a good deal above the true boiling-point, and then boil violently. The boiling-point of the solvent is first taken, then a weighed quantity of substance introduced, and after a few minutes the boiling-point again observed in order to determine the difference.

### § 6. Calorimetry of fluids.

In the changes of volume and temperature which fluids may suffer, heat is generally absorbed, or given out; hence it is necessary to study these changes with the aid of calorimeters, i. e. instruments for measuring quantities of heat. In the light of the principle of conservation of energy, or of the first law of thermodynamics, we may classify heat changes in the following way: when heat is supplied to a fluid—say water—it is spent partly in raising the temperature of the fluid, partly in making it expand; the latter because in order to expand the fluid has to overcome forces and so do work, the heat absorbed being the equivalent of the work so done. The forces overcome, however, are of two kinds: (i) the external pressure to which it is subject (measurable by a pressure gauge); (ii) the internal forces in the fluid. Hence we may write the equation:—

‘Heat supplied to the fluid = increase of heat-content + internal work of expansion of fluid + external work done.’

In terms of the molecular theory these three parts would be explained as being (a) increase in kinetic energy of the molecules constituting the fluid, (b) work done in moving the molecules further apart against the attractive forces existing between them, (c) work done on outside bodies.

The true specific heat of a substance refers to the first term on the right-hand side only, i. e. to the absorption of heat when there is no expansion of the fluid; this is often called the specific heat at constant volume. It is practicable to enclose gases in a rigid vessel so that they may be heated without expansion; if then the heat required to raise 1 gram of the gas through  $1^{\circ}$  be measured



under these circumstances, that is the *true specific heat*; but more frequently the gas is left under constant (e. g. atmospheric) pressure while being heated, and then of course expands. The absorption of heat per gram per degree, in this case, has received the name of *specific heat at constant pressure*; it is greater than the true specific heat on account of the other terms in the equation. In liquids to prevent expansion on rise of temperature would require such enormous pressures as to make the direct determination of the specific heat at constant volume experimentally impracticable; hence it is always the specific heat at constant pressure that is measured; the other can however be calculated from this and other data.

Again, on evaporation of a liquid the volume increases very greatly and a large amount of both internal and external work is done, whilst the temperature remains unchanged. Here, then, only the second and third terms in the above equation have to be considered, and they constitute between them the so-called *latent heat*; the latent heat of evaporation is therefore the heat absorbed by 1 gram of liquid on conversion into saturated vapour, without change of temperature<sup>1</sup>. Evaporation in the lungs and from the skin consequently absorbs a considerable amount of heat from the body; if the air temperature is high, so that the heat carried away by the air is less than usual, the body temperature tends to rise also, but increased perspiration serves to get rid of the excess of heat. The same thing normally occurs when, owing to muscular exercise, the production of heat in the body is above the usual amount. The latent heat of water at the body temperature is 581, i. e. every gram of water evaporated carries off 581 calories (or  $581 \times 4.2 = 2230$  joules of energy).

A similar phenomenon occurs in gases; if a compressed gas be allowed to expand and do external work it needs a supply of heat, if its temperature is to be kept up; this is called the *latent heat of expansion*: if heat be not supplied to it its temperature falls, as in Cailletet's experiment (p. 94).

Quantities of heat are usually measured by the 'water' or 'mixture' calorimeter, i. e. as the name implies, by mixing the

<sup>1</sup> There is a similar latent heat of fusion, on conversion of a solid into liquid; this will be considered in detail below, but meanwhile is needed for the description of one of the calorimetric methods given in this chapter.

substance which is to give out heat with a known mass of water, and observing the rise of temperature of the latter. Thus, to measure the specific heat of a liquid, some 20 or 30 grams of it are heated till the temperature, as read by a thermometer immersed in the liquid, has become steady at some desired point  $t$ , e. g. the temperature of a steam jacket,  $100^\circ$ . Meanwhile a quantity of water is weighed out in the calorimeter. The latter consists of a thin copper or silver vessel, provided when convenient with a lid, and carrying the thermometer, and stirrer. It is placed inside another metallic vessel for protection against loss of heat, and to keep it from contact with the outer vessel is held in place by corks. The temperature of the calorimetric water is then taken ( $t_1$ ), the hot liquid added, and the whole stirred till the temperature is uniform throughout, and the thermometer again read ( $t_2$ ). Then (from the definition of the calorie) the heat received by the water is  $W(t_2 - t_1)$  calories, where  $W$  is the weight of water; this is given out by the substance in cooling from  $t$  to  $t_2$ ; if there be  $w$  grams of the substance and its specific heat be  $s$  it gives out  $w s (t - t_2)$  calories: equating these two quantities we get

$$s = \frac{W(t_2 - t_1)}{w(t - t_2)}$$

To ensure accuracy in the experiment it is necessary, (i) to transfer the hot substance to the calorimeter without loss of heat on the way—this is chiefly a matter of experimental skill; the liquid may be simply poured rapidly into the water, or, if it acts chemically on the latter, may be enclosed in a thin-walled glass bulb and the whole dropped in. (ii) The materials of the calorimeter, thermometer, and stirrer are heated as well as the water, therefore their thermal capacity must be found (from their weight and known specific heat) and added to  $W$  in the above equation in order to get the true thermal capacity of the whole. (iii) As soon as the calorimeter rises in temperature above its surroundings, it begins to lose heat by radiation and conduction; this loss can be allowed for, but it may be made negligibly small by screening off the calorimeter well from its surroundings—its radiation will be least if it be made of polished metal—by using a large mass of liquid and small rise of temperature (with a correspondingly delicate thermometer), and by using water somewhat colder than the air to start with.



To measure the latent heat of evaporation of a liquid the same calorimeter may be employed, with the addition of the apparatus shown in Fig. 40, as designed by Berthelot. *E* is a glass vessel, in which the liquid can be boiled by the ring gas-burner *D*; the calorimeter is protected from being directly heated by the gas-burner by sheets of asbestos. The vapour produced passes down the central tube into the condensing spiral *s*, which is attached to *E* by a ground-glass joint; in this way it has such a short distance to travel from the flask to the calorimeter that it loses

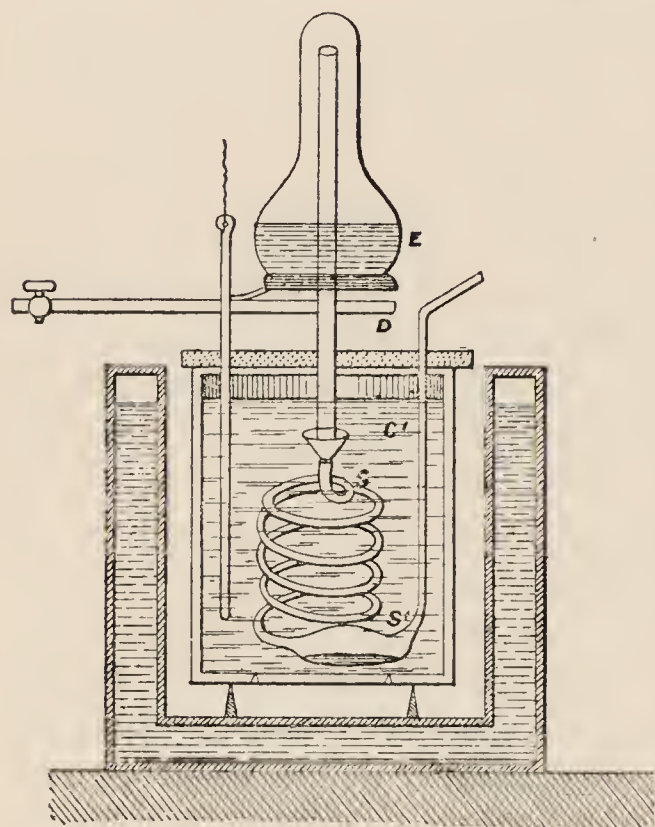


FIG. 40.

no appreciable amount of heat on the way, and at the same time cannot carry over liquid spray with it mechanically. The amount of liquid condensed in the calorimeter is found by weighing, before and after, either the flask, or the condenser *s*: the heat given out by each gram of the vapour can then be calculated. That amount of heat is, however, made up of two portions, (a) the latent heat given out on condensation, without change of temperature; (b) the heat given out by the condensed liquid in cooling from its boiling-point to

the temperature of the calorimeter. The latter quantity must be calculated from a previous knowledge of the specific heat of the liquid, and deducted from the total to arrive at the latent heat of evaporation.

It is not possible to determine the specific heat of gases by the same process as that described above for liquids, because the mass of gas contained in the vessel to be heated would be so very small compared to the mass of the water in the calorimeter. The method adopted by Regnault and others was therefore to cause a continuous stream of gas to flow, at a carefully measured rate, through a spiral tube in a heater, and thence through



a similar spiral in the calorimeter; the quantity of gas used could thus be made as large as required. It is obvious, however, that this method does not give the true specific heat, but the specific heat at constant pressure, since the gas flows without obstruction from heater to calorimeter, and thence out into the air, and is throughout practically at the pressure of the air.

Joly has however succeeded, of late years, in directly measuring the specific heat of gas at constant volume, when sufficiently compressed, by means of a calorimeter of his own invention. A copper sphere, containing the gas compressed into about  $\frac{1}{20}$  of its volume at atmospheric pressure, is suspended by a long fine wire from one arm of a balance, in such a way that it is enclosed in a chamber into which steam can flow; it is balanced with weights, while the chamber is full of air, and the temperature of the air noted, dry steam is then allowed to flow in, the copper sphere and its contents are raised to  $100^{\circ}$ ; but in order to produce this rise of temperature some steam is condensed, and adheres to the metal: the amount thus deposited is determined by the increase in weight recorded by the balance. Now the latent heat of steam at  $100^{\circ}$  is 537 calories, so that if  $w$  grams are deposited, 537  $w$  calories is the amount of heat required to heat the sphere from atmospheric temperature  $t$  to  $100^{\circ}$ , or

$$\frac{537 w}{100 - t} \text{ calories}$$

per degree. This, then, is the thermal capacity of the sphere and its contained gas. By making a similar experiment on the sphere empty, the thermal capacity of the gas alone can be found, and so, dividing by the mass of the gas, its specific heat.

The method can equally well be used to measure the specific heat of solids and liquids.

It is also possible to use the latent heat of fusion of ice for calorimetric purposes. In the earliest attempts to do this the ice melted was weighed, but the process now generally used, due to Bunsen, is to measure the change in volume of the ice on melting. Bunsen's calorimeter is shown in Fig. 4; it is constructed entirely of glass, and consists of the working vessel A filled with air-free water, cut off below by mercury which fills also the side tube B and the capillary C. The water in A is first partly frozen (best by allowing liquid carbon-dioxide and ether to evaporate in the inner tube D) and the whole immersed

in a bath *F* of pure ice; *E* is an air jacket to keep the vessel *A* in better thermal insulation from its surroundings. If then a hot body be dropped into *D* it melts some of the ice in *A*: now, ice is lighter than water, so that fusing some of it causes a contraction; this is measured, after the manner of a thermometer, by the movement of the mercury in the tube *c*, which must be previously calibrated and graduated. The specific volume of water at

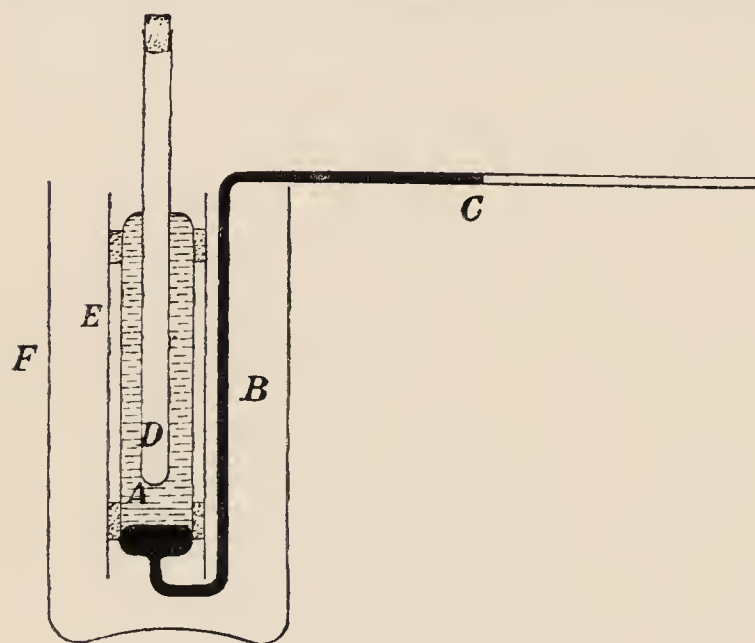


FIG. 41.

$0^{\circ}$  is 1.0001 c.c. per gram, that of ice at  $0^{\circ}$  is 1.0909; hence, when one gram of ice is melted a contraction of 0.0908 c.c. takes place. Now, the latent heat of fusion of ice is 80 calories, i.e. one calorie melts  $\frac{1}{80}$  gram, and consequently causes a change in volume of  $0.0908 \times \frac{1}{80} = 0.001145$  c.c.

Now that electrical measurements can be

carried out with ease and certainty it is becoming common to use electrical heating for calorimetric purposes. Apparatus for measuring the dynamical equivalent of heat electrically has already been referred to (p. 59; see also p. 224); the same apparatus may be employed with other liquids instead of water, and will then serve to measure their specific heats (directly in terms of the mechanical units, or, by comparison with the experiment made on water, in terms of the ordinary calorie).

The following table will serve to give a notion of the results obtained:—

	Specific heat.	Latent heat of evaporation.
Water	1.00	536.
Alcohol	0.61	206.
Benzene	0.38	94.4

It will be noticed what a very large ratio the latent heat bears to the specific heat. The latent heat is in a sense an expression of the difference in character between the liquid and saturated vapour, and, as we have seen that that difference is reduced with

rise of temperature, till finally at the critical point it vanishes and the two forms of matter become identical, it follows that the latent heat of evaporation must also diminish with rise of temperature, and become zero at the critical point. The usual forms of apparatus do not admit of measuring the latent heat when a liquid is boiled under a pressure greater or less than that of the atmosphere, but special experiments have shown the above conclusion to be true.

The latent heat of evaporation consists of two parts, the work done in overcoming the internal forces (cohesion) of the liquid when its parts become separated so as to occupy the much larger volume of the vapour, and the work done in overcoming the atmospheric pressure during evaporation. The latter part is easily calculated; the vapour behaves nearly as a perfect gas, so we may calculate its volume from the characteristic equation  $v = \frac{RT}{p}$ ; the volume of the liquid is so small that it may be neglected by comparison; hence we may say that when one gram-molecule of the liquid evaporates its volume increases by  $\frac{RT}{p}$ , where  $R$ , the gas constant, refers to the molecular weight (p. 79), and the evaporation takes place under the pressure  $p$ ; hence the work done against that pressure = pressure  $\times$  increase in volume =  $RT$ . Now  $R = 2$  calories nearly, so that the thermal equivalent of the external work done is simply  $2T$  calories per molecule, where  $T$  is the absolute temperature at which the evaporation takes place. Choosing benzene as an example, the molecular weight ( $C_6H_6$ ) being 78, the molecular latent heat is  $78 \times 94.4 = 7363$  cal. at the boiling point  $80^\circ$  cent. Then  $T = 273 + 80 = 353$ , and the external work  $2 \times 353 = 706$  cal., or about one-tenth of the whole. The remainder  $7363 - 706 = 6657$  cal. is called the *internal latent heat* (per mol.).

The specific heat of a gas is also most conveniently expressed per gram-molecule. The following numbers are due to Regnault :—

	Under constant pressure.			Mol. sp. heat at constant volume.
	Mol. wt.	Sp. heat.	Mol. sp. heat.	
Hydrogen . . . .	2	3.41	6.82	4.84
Nitrogen . . . .	28	0.244	6.83	4.85
Oxygen . . . .	32	0.218	6.98	5.00
Carbon dioxide . .	44	0.216	9.50	7.52



The last column is derived from that before it in the following way. When a gram-molecule of gas at  $T$  degrees (absolute) and pressure  $p$  is heated through  $1^\circ$  at constant pressure, its volume, which was to start with  $= \frac{RT}{p}$ , becomes  $\frac{R(T+1)}{p}$ , the increase amounting to  $\frac{R}{p}$ . If this be multiplied by the pressure  $p$  under which the expansion takes place we get the work done against the external pressure  $= R$ , or about 1.98 calories. This is independent of any work done against the internal, molecular forces in the gas; but there is hardly any such work in a gas, so that when the gas is heated under constant pressure the absorption of heat per degree is greater than the true specific heat (or specific heat at constant volume) by just 1.98 calories.

The point assumed in the above argument, viz. that the internal forces in a gas are so small that the work done against them may be neglected, was the subject of direct experiment by Joule. Two copper vessels, one containing air at high pressure, the other evacuated, were immersed in a calorimeter; a tap between them was then opened and the pressures equalized. Here no work is done on any outside bodies, so that if any heat was absorbed by the expanding air, it must have been to do work against the internal forces in the air; no change in the temperature was noted, however, so it was concluded that no work was done. The experiment was not very delicate, and subsequent experiments by a different method, performed by Joule and Thomson (Lord Kelvin) together, showed that in unresisted expansion air is slightly cooled, and hydrogen slightly warmed: there is consequently a slight effect due to attractions and repulsions between the molecules of gases. The influence of this on the specific heat need hardly be taken into account, but the cooling effect in air has nevertheless been made use of by Linde for the liquefaction of air, as already mentioned (p. 94).

## § 7. Surface tension

The surface layer of a liquid possesses certain properties that are not shared by the mass of it, for in the interior the arrangement of matter is symmetrical round any point, whereas for a point on the surface the surroundings consist of liquid on one side only, while on the other is solid or gas, or perhaps a different liquid. The molecular theory throws light on this point, for whilst in a gas the molecules practically free from one another's influence fly about freely with high velocities, and so produce the phenomenon of pressure on the walls of the containing vessel, in a liquid the mutual attractions of the molecules are great enough to keep the substance together in a definite volume; we have in fact just seen that to separate the molecules of a liquid

till they are free from one another, and so convert it into vapour, requires a very large amount of energy—the so-called latent heat of evaporation. The molecular attractions in a liquid are thus very great, so that a molecule of the surface layer is strongly pulled inwards, while there is no corresponding force exerted by molecules outside it (if there be gas on the other side of the liquid surface, a different force if there be solid or another liquid). This is just as if the surface layer of liquid constituted a stretched elastic skin; it produces a pressure inside it and is itself in tension—has a tendency to rupture across any line drawn in the surface.

We might imagine a mass of liquid drawn out into a thinner and thinner film until there was only one molecule in the thickness of it; in this way the molecules would eventually all be separated from one another, and the liquid converted into gas. It is probable that the film would actually break before it reached that degree of thinness, but the imaginary process enables one to see that the work done in drawing out such a film would be identical with that done in evaporating the liquid in the ordinary manner, and even to calculate roughly, as Lord Kelvin has done, the distance between the molecules in the liquid. The same argument indicates that there should be a certain parallelism between the latent heat of evaporation and the surface tension of a liquid. This is in fact the case; water, e. g., has a very high value for each. Further, the surface tension, like the latent heat, should fall off with increase in temperature, and vanish at the critical point where the distinction between liquid and vapour vanishes; this also has been proved experimentally and the surface tension found to fall off nearly uniformly as the critical temperature is approached.

The effect of surface tension is most simply seen in the case of a free drop of liquid: a rain-drop is spherical in shape, and so is a drop of oil immersed in a mixture of alcohol and water of the same density, so that it may float freely without any tendency to rise or sink. In these cases there is nothing to prevent the tension in the surface layer of the liquid causing it to contract as much as possible; now, a given volume of liquid possesses least surface when its shape is spherical, therefore if there are no other forces than the surface tension acting it will assume that shape. If, however, a drop of liquid which does not wet glass—



say mercury—be placed on a glass plate its weight will tend to make it spread out flat, while the surface tension tends to make it spherical, so that the actual shape will be between the two; a very small drop will be nearly spherical, but a very large one, on account of its greater weight, will form a flat disc with rounded edges.

A stream of water in descending (e.g. a thin stream from a water-tap) acquires more and more speed, and consequently becomes thinner and thinner as it descends; but, owing to its being stretched out in this way, the amount of surface for a given volume of liquid increases till it eventually breaks up into drops, so that by the rearrangement of shape the water may come to have less surface than if it remained in a stream. In the same way if liquid flows out from a fine orifice, such as the end of a capillary tube, it breaks up at once into drops; if a drop be watched (coming out of a burette with the tap nearly turned off) it will be seen to grow to a fixed size before breaking away, the surface tension acting all round the rim of the orifice holds it up until the weight becomes too great. Thus, in a drop of water coming out of a hole 0.2 cm. in diameter, the tension is acting round the circumference, which is about 0.6 cm. long; the surface tension of water is about 75 dynes per cm., so that the total force supporting the drop is in this case  $75 \times 0.6 = 45$  dynes. Now the weight of a gram is 981 dynes (p. 17), therefore the drop will break off when it contains  $\frac{45}{981} = 0.046$  grams of water; the weight of a drop falling from any given tube is thus definite in amount.

Another simple case is that of a bubble. A soap-bubble—an obvious instance of the tendency of a liquid to assume a spherical shape in the absence of other forces—has two surface layers, one inside and one out: both these tend to produce a pressure on their concave side, i. e. the inside of the bubble. The pressure is in inverse proportion to the diameter of the bubble, and for a soap-bubble of half a centimetre diameter it amounts to about the equivalent of one millimetre of mercury; hence, when a bubble breaks (owing to the film becoming too thin in some spot) it does so explosively, and projects its own substance to quite considerable distances.

When a liquid is boiled small bubbles form inside it, and the liquid round these evaporates into them, and the bubbles thus



grow in size as they rise to the surface. If there be a good supply of air-bubbles in the liquid it boils easily and quietly; this can be accomplished by leading a fine stream of air in by a capillary tube, or by materials like pumice-stone, beads, or bits of platinum foil, which carry a little air with them. But if there is no such supply of air-bubbles boiling becomes difficult, for, as we have just seen, the gas inside a bubble is at a higher pressure than that in free space outside, and the more so the smaller the bubble; hence, when only extremely minute bubbles are present to start with, the vapour must reach a higher pressure than the atmospheric in order to form inside them, and the liquid therefore gets raised above the true boiling-point. In water very carefully freed from air this superheating may amount to many degrees, and then when boiling does take place it is almost explosive in character.

It may be noted in this connexion that condensation of water vapour in air similarly needs to be started by the presence of 'nuclei'—dust particles in this case. Air that is freed from dust by filtration through cotton wool can be considerably oversaturated with water-vapour without condensation; if then a little dust be mixed with it a cloud at once forms, consisting of a small drop of water condensed on each dust particle.

A liquid placed on a solid surface either stands in a drop or spreads out indefinitely into a film: the latter is the case for most liquids in contact with clean glass or metal surfaces; and when it is so it results in a tendency to raise the liquid against gravity. Thus, if a small tube be dipped in such a liquid the tension in the film will lift the liquid to a certain height in the tube; if  $a$  is the radius of the tube there is a tension  $T$  acting all round the circumference which is  $2\pi a$  long: hence, if the density of the liquid is  $d$  and it rise to the height  $h$  there is a volume of it  $= \pi a^2 h$  raised; this has a mass  $\pi a^2 h d$ , and weighs  $\pi a^2 h d g$  dynes, so that  $2\pi a T = \pi a^2 h d g$ , or

$$h = \frac{2T}{a d g}.$$

Thus, in a tube of  $\frac{1}{2}$  mm. radius water with a surface

tension of 75 dynes per cm. will rise to a height  $= \frac{2 \times 75}{0.05 \times 1 \times 981}$

$= 3.06$  cm. The best way of measuring  $T$  is to observe the rise in a tube of measured radius. It will be noted that the height to which the liquid rises is inversely proportional to the diameter

of the tube hence the rise is great in a so-called 'capillary,' or tube of hair-like fineness.

Liquids show the same tendency to rise in narrow spaces of other shapes—as between two glass plates—and even merely against the face of a solid immersed in them; as may be noticed with water standing in a glass vessel, the edge of the water is always curved upwards. Still more marked is the tendency to rise up porous materials, such as cloth, sugar, porous earthenware and the like. The liquid thus drawn up in films and exposed to air tends of course to evaporate, so that in course of time quite a considerable quantity may be lost in this way, and if the liquid be a mixture or solution the more volatile part will evaporate more quickly, leaving the others behind. This accounts for the creeping observed with salt solutions in the necks of bottles and similar places; the water evaporates from the film, leaving a crust of salt behind.

Surface tension doubtless helps to cause the rise of sap in the fine capillary tubes of plants, when their roots are immersed in water or damp earth, but it is not the sole or even most important cause (see osmosis, p. 128).

Water may always be prevented from spreading over a surface by making the latter greasy, because grease and oils have a lower surface tension than water and aqueous solutions. In the same way it is the very great surface tension of mercury that prevents it spreading out into a film. Mercury, therefore, behaves oppositely to water in capillary tubes and the like: it stands lower in a tube than in a wide vessel, and the lower the narrower the tube; this may cause considerable inaccuracy in a mercury gauge, such as a barometer. A good barometer should be at least 1 cm. wide at the top, where the mercury surface is, and a U-gauge of mercury is also erroneous on account of surface tension, for, as may easily be noticed, the mercury is more convex when the column is rising, and may become nearly flat when falling; this causes a difference between the effect of the surface tension on the two sides. It may be minimized by shaking or tapping, but when the gauge is used for rapid fluctuations, as in a sphygmograph, causes unavoidable error. A U-gauge of oil, water, or other liquid that wets glass, is not exposed to this source of error, because the film over the glass ensures that the curve of the liquid is the same on both sides, and consequently the effects of



the surface tension on the two sides neutralize. Mercury in contact with any surface tends to be depressed ; and when a solid is lowered into mercury, since the mercury surface has to be increased in area by the process, its tension offers a resistance, so that, apart from the great buoyancy of the liquid metal, it is difficult to push even a sheet of paper into it.

### § 8. Flow of liquids.

When a liquid is contained in a vessel it presses, as we have seen, against the solid walls of the vessel in all directions, and to the same extent at all points which are at the same depth below the surface. The actual pressure exerted on a square centimetre is  $h d g$ , where  $h$  is the height of the free surface of the liquid,  $d$  its density, and  $g$  the acceleration of gravity (981). If now an opening be made in the vessel, the liquid will flow out, and the rate at which it flows depends on the pressure of the liquid. We may find the exact relation by considering the amount of energy that the liquid in the vessel possesses : suppose a gram of liquid to flow out, then practically it comes from the top, since it is there that the liquid disappears from inside the vessel. We may then consider that the gram in question has fallen through a height, or 'head,'  $h$  (in centimetres) from the free surface to the opening ; in doing so it gives out  $h g$  ergs of work (p. 22). Now the kinetic energy of a body moving with velocity  $v$  is  $\frac{1}{2} \times \text{mass} \times v^2$ , or for one gram  $\frac{1}{2} v^2$  ; hence, if all the work that the liquid can do is spent in giving kinetic energy to it, the velocity with which it will flow out is given by setting the kinetic energy = work done, or—

$$\frac{1}{2} v^2 = g h, \text{ whence } v = \sqrt{2 g h} \text{ or } h = \frac{v^2}{2g}.$$

The velocity can in no case exceed this, and will fall short of it, if there be any frictional resistance to absorb part of the work done.

A liquid has not necessarily a free surface, but may be completely enclosed in solid vessels, as is the blood in the circulatory system, or the water in a system of hydraulic pressure mains. The pressure in the blood is mainly caused by the tension of the walls of the vessels, so that if a long open tube of small bore were inserted at one end into a blood-vessel, while the other was held vertically, the blood would rise in it to a point much



above the highest point of the body ; it would form a free surface in the long tube, and the head  $h$  in the above calculation must be reckoned from this (imaginary) free surface downwards : it is naturally greater for a point low down in the body than for one high up. Thus the average blood pressure in a rabbit may be such as to raise a column of blood in the inserted tube about 190 cms. high ; consequently if a cut be made in an artery of a rabbit the speed with which the blood will be ejected is found by putting  $h = 190$  in the above formula, or—

$$v = \sqrt{2 \times 981 \times 190} = 611 \text{ cms. per sec.}$$

When a liquid flows out through an orifice in the side of a thin-walled vessel, the streams coming from various parts and converging to the orifice continue to converge after passing it for a certain distance, so that the proper section of the issuing stream is only obtained (at a distance of something like the diameter of the hole) when the convergence has ceased, and all the liquid is flowing in a parallel stream ; its area is then only some 64 % of the area of the orifice, so that in calculating the amount of liquid flowing out, allowance for this ‘contracted vein’ must be made. This, however, represents no loss in energy but only in the quantity flowing, and to avoid it it is customary to provide a nozzle for the liquid to flow from shaped to the jet. There is, however, also a frictional resistance to the outflow, which in case of a hole in a thin wall may reduce the velocity from 1 to 4 % ; this means a decrease both in the quantity flowing out and in the energy the liquid possesses : it may be expressed by saying that part of the head of liquid is used up in friction, leaving the head available for giving velocity to the water (or so-called velocity head  $h = \frac{v^2}{2g}$ ) effectively less. When the liquid has to flow through a long pipe (such as a water main) much more friction occurs, the head lost in overcoming it being about proportional to the length and inversely to the diameter of the pipe, e. g. in a pipe that was 40 times as long as it was broad friction might use up about half the head. This friction, however, is mainly that occurring between the liquid and the solid surface of the pipe (skin friction). In narrow tubes the conditions of flow are different ; here the layer of liquid in immediate contact with the wall of the tube is at rest, so that there is no

skin friction, but the various layers of liquid have to slide over one another, the velocity increasing steadily as one passes from the wall to the centre of the tube; consequently the internal friction or viscosity of the liquid comes into play. It is to the class of 'narrow' tubes in which viscosity has to be considered that blood-vessels almost exclusively belong.

In such narrow tubes the resistance to the flow dependent on the viscosity of the liquid may be shown to be proportioned to the length of the tube, and, inversely, as the fourth power of its diameter—a much greater rate of variation than holds for the flow in wide pipes: e. g. by reducing the diameter of the tube to one-half the volume of liquid flowing through it is reduced to one-sixteenth; if the diameter be one-third, the flow is only  $\frac{1}{81}$ .

The mode of dependence on the length may best be expressed, as in the corresponding case of the flow of heat (p. 70), by the conception of a slope or gradient. Fig. 42 illustrates this: it represents a model, such as could easily be made of glass tubing, showing a horizontal tube (wide or narrow) AB, into which are inserted at intervals narrow vertical gauge tubes C, D, E, to indicate the head at various points. If the tube AB be of uniform bore, so as to offer an

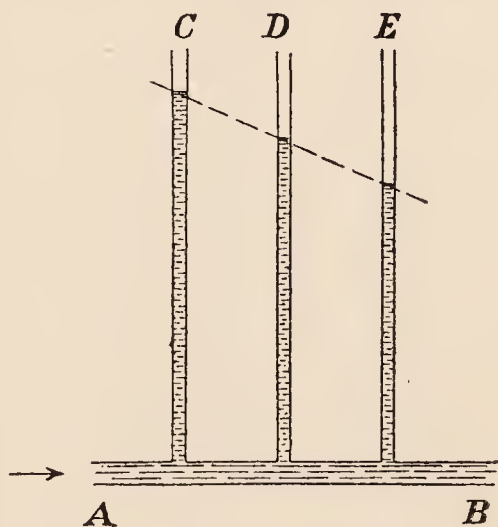


FIG. 42.

equal resistance to the flow of liquid at all points, it will be found that the liquid surfaces in C, D, E, lie in a straight line, sloping downwards from the end at which the water flows in to that at which it flows out. Thus at C the pressure of liquid is that indicated by the column in the gauge at C; going to D a part of that pressure has already been used up in overcoming the friction in the tube between C and D, so that what remains is measured by the shorter column at D; again, going an equal distance from D to E the pressure is again reduced by an equal amount. The rate at which the pressure changes, or

$$\frac{\text{difference of pressure between A and B}}{\text{distance A to B}},$$

is called the pressure gradient between those points; and the

pressure gradient it is that causes a flow of liquid in the pipe in just the same way as a temperature gradient causes a flow of heat through a plate (p. 70).

In the case of narrow tubes, in which only the viscosity has to be taken into account, the flow of liquid is proportional to the pressure gradient producing it. This is not true of wide pipes however. The exact relation for narrow tubes is—

$$\text{Volume flowing per sec.} = \frac{\pi}{8} \frac{r^4}{\eta} p$$

where  $r$  = radius of the tube,  $p$  = pressure gradient,  $\eta$  = coefficient of viscosity, a constant expressing the specific properties of each liquid (or gas):  $\eta$  has the following values for some important fluids—

Water at	0°	0.0181	Mercury at	10°	0.0162
„	20°	0.0102	Ether	„ 10°	0.00283
„	90°	0.0032	Air	„ 10°	0.000172

(lengths in cm., pressure in  $\frac{\text{dynes}}{\text{sq. cm.}}$ ).

From the numbers in this table it may be noted, first, how much viscosity is diminished by rise of temperature. (This is illustrated by the fact that filtration can be effected far more quickly hot than cold; the liquid takes less time to flow through the fine meshes of the filter paper.) Secondly, that air is about a hundred times less viscous than water, so that with the same driving pressure a hundred times as great a volume of air will flow through a capillary tube as of water.

When liquid flows through a channel of varying size, as the liquid is practically incompressible, the volume flowing per second across any section of the tube is necessarily the same, and consequently the rate at which it flows is inversely proportional to the cross section. But, as we have seen, a stream of liquid possesses kinetic energy proportional to the square of its velocity, or in

other words a velocity head =  $\frac{v^2}{2g}$ ; consequently the velocity

head is greatest in the narrow parts of the tube. Now at any point the total head of the liquid, expressing the total amount of energy it has, is divided between velocity head and the actual head which is registered by a gauge there, and indicates the potential energy of the liquid; hence, if the velocity head be increased, the actual head or pressure must be decreased by an equal amount. Fig. 43 is intended to represent this. The pressure as read by the gauge D attached to the narrow part of the tube is less than the mean between those at C and E,



because at D more of the energy of the liquid is absorbed in the form of kinetic energy, so that less is available to produce pressure. The phenomenon is of course complicated by the gradual change in pressure along the tube due to friction; in the figure the pressure at D is represented as actually less than that at E further on: this will be the case if the constriction be sufficiently abrupt. On the other hand, it is clear that if there be an expansion in the tube the pressure will be greater than without it. The action of the common filter pump (p. 45) depends on these facts: a stream of water is made to flow through a very small jet, its velocity rises so high there as to account for nearly all the energy of the water, so that the pressure at the jet is actually below that of the atmosphere; hence, if the jet be surrounded by a small chamber with a side tube, air is sucked into it and carried away by the flow of water.

Where, however, as in the circulation of the blood, the velocity is in any case low, the effects due to changes in velocity head are not important, and it is of more consequence to pay attention to the effect of variation in bore of the tubes on the frictional resistance. If the same stream has to flow through a wide tube and a

narrow one consecutively (the two tubes are said to be *in series*) the pressure spent is in proportion to the resistance overcome: this is much greater in the narrow tube than in the wide one, so that a much greater drop in pressure will occur for each centimetre length in the narrow tube than in the wide one—this is the case, e. g., in a system of water mains; if one tap only be turned on there may be as much pressure lost in the few metres of narrow pipe leading to it as in a much longer length of the street main supplying it. Still more is that the case when the tubes are capillary, since, as we have seen, the flow then varies inversely as the fourth power of the diameter when the pressure gradient is kept the same; or, in other words, if the flow of liquid through the tubes is the same, the pressure gradient must be directly proportional to the fourth power of the diameters. To illustrate

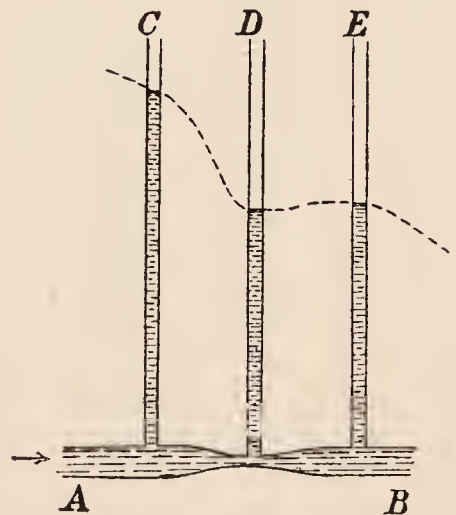


FIG. 43.

this and the following case, let us suppose that a head of 2,000 mm. drives a current through two tubes in series, a wide one (representing an artery) and a narrow (arteriole), and that the latter has 2,000 times the resistance of the former : then (leaving velocity head out of account)  $\frac{1}{2001}$  of the head will be spent in over-

coming friction in the artery,  $\frac{2000}{2001}$  in the arteriole, or say 1 mm.

and 1999 mm. nearly, so that a gauge inserted in the far end of the artery will only register 1 mm. less pressure than one in the near end.

Now, suppose that two tubes are set side by side, so that the liquid can flow through either (the tubes are said to be *in parallel*), the same driving pressure being applied to the pair as originally to a single one ; then the flow through each will be independent of the other, and if the two tubes be alike, the total flow will be just double that which either would give alone. This, however, is not the actual arrangement of water pipes or blood-vessels ; the two would be branches of one larger tube. In this case let us again suppose that each small tube has 2,000 times the resistance of the large one, then the two tubes in parallel will have only 1,000 times that resistance, since each of them has only to carry half the flow. Hence  $\frac{1}{1001}$  of the pressure will be required in the large tube,  $\frac{1000}{1001}$  in the small ones, or approximately 2 mm. and 1998 mm. respectively ; i. e. the pressure on each small tube will be a trifle less than when there is only one, and the total flow consequently a trifle less than double that through a single small tube. If there were a hundred small tubes, the resistance of them would be  $2000 \div 100 = 20$  times that of the large one, and  $\frac{1}{21}$  of the total pressure, or nearly 100 mm., would be spent in driving the flow through the large tube.

In the circulatory system, when an artery divides, the branches always have more cross section between them than the original vessel, but, on account of the great influence of the diameter on friction, the branches have more resistance between them than the wide vessel. Hence (i) in the capillaries the rate of flow is slower than in the arteries and veins ; (ii) most of the driving pressure is spent in the capillaries ; (iii) if one part of the capillary system is choked, and the flow through it diminished, this throws a greater pressure on to the other parts, and their supply is conse-



quently increased, unless the pressure produced by the heart is simultaneously lessened ; (iv) on account of the lower velocity in the capillaries, and consequent lower velocity head, the pressure there is greater than the mean between that in the arteries and veins.

The actual flow of blood is considerably modified by the fact that the pressure driving it is not constant, as is assumed in the foregoing discussion, but is due to the intermittent action of the heart. We have not space here for a complete discussion of the somewhat complicated phenomena resulting from this, but certain leading points may be mentioned. When an inflow of liquid into a tube takes place, if the walls be very rigid—say of metal—the pressure causes an imperceptible expansion of the tube, and consequently the same quantity of liquid must flow out at the far end almost simultaneously ; but if the walls yield easily, such as those of an artery in its normal condition, the immediate effect is to cause the near end to expand. The elastic reaction of the walls causes them to contract again, and in doing so to force the excess of liquid into the next section of the artery, and make it expand ; in fact a *wave of expansion*, or *pulse*, is set up, and propagated with a definite velocity (p. 35), and the outflow at the far end does not take place till the wave has had time to reach there. A single impulse of this kind would cause the artery to go through a series of expansions and contractions, for the same reasons that a pendulum swings to and fro ; but a rapid succession of pulses is produced by the heart, the consequence being that the artery has not time to contract entirely after one, before the next makes it expand again. Hence it is always in a state of tension, although to a fluctuating amount.

Now the effect of friction is to attenuate the fluctuations, so that at the far end of the artery, though the average pressure of the blood is not much less than at the near end, the actual value of it varies between much narrower limits. Still more is this the case in the smaller vessels ; indeed, the flow through them is practically uniform. Hence, if an artery be cut the blood escapes intermittently, but from a vein the flow is continuous. If the walls of the artery are abnormally rigid, the wave of pressure (pulse) is transmitted much more rapidly (this does not affect the rate of *flow* of the blood), and the effect of friction is less marked, so that the pulse at the far end of the artery is abnormally violent.



## § 9. Diffusion : osmosis.

The movements considered in the preceding section were movements of fluids in mass ; even in the case of flow through a capillary tube (sometimes called 'transpiration') there is a visible stream of fluid, so that if a solid particle be placed in it, it will be carried along by the stream. Other movements are possible, however, on account of the mobility of the molecules of a liquid or gas, and are known under the general name of *diffusion*. Thus, if a jar full of hydrogen be exposed mouth downwards to the air, it is found after a little while that there is no hydrogen left in it ; the mouth being at the bottom the hydrogen cannot have flowed out, for it is lighter than air, and indeed if the most delicate vane or indicator be placed in the neck of the jar it will not indicate any current. The movement has taken place molecule by molecule ; neither the outflow of hydrogen nor the inflow of air constitutes a stream of those gases anywhere, but a diffusion of them, for the space in which the diffusion takes place everywhere contains hydrogen and air mixed. The mechanism may be illustrated by the difference between the movement of a thousand persons in the street threading their way east and west amongst each other individually, and the movement of five hundred soldiers marching east simultaneously with the same number marching west. Since we are unable to distinguish individual molecules, we are unable to see the mechanism of processes of diffusion, although they may take place rapidly.

Diffusion of fluids takes place according to laws similar to those of the diffusion (or conduction) of heat (p. 68). The driving force here is difference in concentration of the diffusing substance in different places. Thus, in the jar of hydrogen the concentration of the hydrogen—amount in unit volume—is large, in the space outside it is zero ; there exists, then, a *gradient of concentration* between those two places, and this will cause the hydrogen to flow from where its concentration is high to where it is low. If there were no air present the flow would be extremely rapid ; the concentration, and consequently the pressure, of the gas would become equalized in a second or two. The presence of air tending to diffuse in the opposite direction (since the concentration of the air is larger outside than inside the jar) acts as

a great hindrance ; with both gases present the total pressure of the gas is the same inside and outside the jar, but the *partial pressures* of the hydrogen and air respectively are different, so that the process is one tending to equalize the partial pressure of each gas throughout the space open to it. We may therefore equally well look upon the difference of partial pressure as the effective cause of diffusion, and refer to the gradient of pressure. The rate of diffusion between each pair of gases depends on their specific properties, so that the general conclusion may be put in the form—

‘Amount diffusing across 1 sq. cm. of a surface per sec.  
= diffusivity  $\times$  gradient of concentration across the surface.’

It is by diffusion that gases are exchanged in the lungs: the so-called tidal air of the lungs, that which is drawn in and out by the movement of the muscles, constitutes less than half of the total air contained. The carbon dioxide given off from the venous blood is consequently not directly expelled from the lungs, but remains in the residual air ; from this it diffuses into the tidal portion and so is expelled, for the concentration of carbon dioxide in the residual air in contact with the epithelium is constantly greater than that in the external atmosphere. Similarly, the concentration of oxygen in the atmosphere is greater than in the interior of the lungs, so that it constantly diffuses inwards, and is removed by the blood that absorbs it. So, too, carbon dioxide is absorbed by the tissues below the stomata of plants ; its concentration is consequently less there than in the air outside, and so diffusion occurs, bringing a constant fresh supply of carbon dioxide for assimilation.

Again, evaporation is always associated with diffusion : if a dish of water is exposed to the air evaporation takes place at its surface, and the layer of air in contact with it becomes saturated, or nearly so ; but the concentration of the aqueous vapour in it being then greater than elsewhere it diffuses away, and more water evaporates to take its place.

In all such cases movement of the gases in mass, or convection, assists diffusion, as it does in the case of heat conduction, for it brings into close contact layers of gas of very different composition, and so increases enormously the gradient of concentration ; thus, the simple process of diffusion would be insufficient to carry on respiration in the higher animals, and the respiratory movements are made in order to increase the facility for diffu-



sion. So, too, diffusion (through windows) may be insufficient to ventilate a room satisfactorily, i. e. remove the products of respiration and restore the normal percentage of oxygen in the air; but if a wind (i. e. movement of air in mass)—artificial or natural—assist it the ventilation is largely increased. So, too, a liquid can be made to evaporate far more rapidly by blowing a current of air over it.

Liquids diffuse in the same manner as gases, though more slowly; but whereas any two gases are miscible, and so diffuse into one another when placed in contact, the miscibility of liquids is, as we have seen (p. 98). more restricted. We may take, however, two liquids capable of mixing—such as alcohol and water, and place the lighter—alcohol—over the other in a beaker, pouring it on very cautiously, so that no streaming of the one into the other takes place; but diffusion will at once begin, and in course of time a uniform mixture of alcohol and water will be found throughout the vessel. The diffusion takes place according to the same laws as in gases, but if convection currents be completely avoided it will be weeks before the composition is sensibly the same throughout, although by stirring the diffusion is so enormously facilitated that the same state may be reached in a few seconds.

Substances in solution diffuse similarly, and two aqueous solutions are practically always miscible.

The rate of diffusion, both of gases and liquids, varies according to their specific nature—a fact which is taken account of in the equation on p. 121, by means of the factor there called the diffusivity; generally, the lighter the molecules (smaller the molecular weight) of a substance the more rapidly does it diffuse. Thus, hydrogen diffuses much more rapidly than other gases, and in solution light molecules, such as those of salts, diffuse more rapidly than larger molecules of organic substances, such as sugar ( $C_{12}H_{22}O_{11} = 342$ ); whilst the bodies with the largest molecular weight (5,000 and more) are the proteids, and these are hardly capable of diffusion at all. It may be noted that some membranes of the body do not behave indifferently towards all dissolved substances, but exercise a selective—probably chemical—action. Thus, it has been observed that dextrose diffuses through the intestinal membrane faster than sodium sulphate, despite the high molecular weight of the former.



The dependence of diffusivity on molecular weight is in accordance with the movements of fluids in other ways—flow through capillary tubes, and flow from an orifice; considering the latter only, the velocity is  $\sqrt{2gh}$ , where  $h$  is the head of the fluid. Now, if two gases, say oxygen and hydrogen, are at the same pressure, to express the pressure as a 'head' of gas, that of the hydrogen must be sixteen times the greater, since its density is sixteen times less; consequently hydrogen will flow out faster than oxygen at the same pressure, in the ratio of  $\sqrt{16} : 1$ , or  $4 : 1$ ; and the rate of diffusion is also approximately four times as great for hydrogen as for oxygen. This is explained on the molecular theory, by the equality between the average kinetic energy of translation of the molecules of any two gases at the same temperature. But the kinetic energy is measured by  $\frac{1}{2} \text{ mass} \times (\text{velocity})^2$ , so that as the mass of an oxygen molecule is sixteen times greater than that of a hydrogen molecule, its velocity must be four times smaller in order that the two should possess the same energy: this greater molecular velocity of the hydrogen may be taken as the cause of the greater velocity of outflow, of transpiration, and of diffusion.

Diffusion, like transpiration, is considerably accelerated by rise of temperature.

Diffusion can take place, not only when there are differences in composition in a fluid that is continuous, but in very many cases also between two fluids separated by a solid partition. In the case of gases, such materials as porous earthenware and plaster of Paris are *permeable*, and so allow diffusion to take place through them. If a porous pot (such as that used for Daniell cells) be fitted air-tight with a rubber stopper carrying a glass tube, and be placed with the tube downwards in a beaker of water, then, on surrounding the jar with a large inverted beaker of hydrogen, bubbles will be driven out from the glass tube through the water. Hydrogen has diffused into the jar, and air that was originally in it has diffused out; but as the movement of the hydrogen is the more rapid there is at first an excess of pressure in the jar: if, however, the apparatus be left long enough, care being taken that the hydrogen atmosphere round the jar is maintained, the pressure will eventually be the same inside the jar as outside, and that, whatever the gas used in the experiment may be. So that the different rates at which the gases pass through the porous material depend only on the nature of the gas; the solid allows them all to pass indifferently, and is simply permeable.

Diffusion takes place, similarly, in the case of liquids and solutions. Porous earthenware, paper, and especially the membranes of animal and vegetable tissues show this phenomenon ; it is by this means that food substances, after solution in the alimentary canal, pass into the blood and thence into the various tissues of the body. Moreover, dissolved gases also pass freely by diffusion, both through a membrane, and out of or into the solution, from the atmosphere : thus, if a glass of water saturated with carbon dioxide (such as soda-water) be left in the air, in course of time nearly all the carbon dioxide will disappear from it (not quite all, because there is some in the atmosphere), and oxygen and nitrogen will diffuse into the water till a state of equilibrium is set up. If the water be covered with a permeable membrane the same process will go on, only somewhat more slowly, and this is what takes place in respiration. Fishes breathe by diffusion, through the walls of their gills, of the carbon dioxide produced in the body outwards into water, and of the oxygen of water inwards, on account of the deficiency of that gas in the venous blood. Land animals similarly effect a direct exchange of gases between the venous blood and the air through the membrane of the lungs.

Passage through a porous partition always involves overcoming a certain resistance, on account of the extremely narrow channels offered to the diffusing molecules, so that the diffusion is slower than without it ; thus, e. g., a porous jar is used, in making up a Daniell cell (p. 192), to keep the solutions of zinc and copper sulphate from mixing rapidly. If the partition be permeable for all the materials present in the fluid, then, eventually, the composition of the fluid will be the same on both sides ; but, as we have seen, the rate at which diffusion takes place is very different for different substances, so that temporary differences of composition may easily be set up by diffusion, as in the experiment with hydrogen mentioned above. Thus, in the respiration of fishes, if their blood remained in contact with the outside water for long, solids dissolved in it would diffuse out through the branchial membrane, as well as gas ; but, in the short time that the blood is exposed to diffusion, only the more quickly diffusing gases are got rid of in any large quantity, the solid dissolved constituents of the blood being retained.

Further, many membranes are not indifferently permeable to



all substances, but exercise a selective action on them. Thus, if a membrane separates pure water from a saturated salt solution, water will flow through it to the side of the solution, and salt to the side of the pure water : the weight of water passing through in an hour will be greater than that of salt, but how many times greater depends on the properties of the membrane to a certain extent. Generally speaking, narrowing the pores of the membrane hinders the passage of salt more than that of water. The ratio of the quantity of water to that of salt passing has sometimes been called the endosmotic equivalent.

The selective action of the partition may, however, go further than this, and prevent entirely the passage of one substance, while allowing that of the other to take place with comparative freedom. Such a partition is called *semi-permeable* : its properties are different in important respects from those of an indifferently permeable one, for a permanent difference in composition will exist between the fluid on the two sides of it.

Such cases are known among gases : e. g. palladium is permeable to hydrogen, but not to other gases. If, then, a glass tube be closed at one end by a sheet of palladium foil, whilst the other end is connected to a pressure gauge, we shall be able to study, by means of the pressure, the behaviour of the palladium to gases. First, if the tube be filled with hydrogen, and be enclosed in a vessel full of hydrogen, the gas pressure inside and outside the tube will, after a time, be found equal, showing that the partition is permeable to that gas. If the experiment be repeated with nitrogen it will be found that the outside and inside pressures have no influence on each other ; if they are unequal to start with, they will remain unequal, for the palladium does not let any nitrogen through. Next, let nitrogen be put inside the tube, and its pressure be measured, then enclose the tube in a vessel of hydrogen and measure the pressure of the latter. After a sufficient time the hydrogen will be found to have penetrated the tube, just as if nitrogen were not there : its *partial pressure is equal to the pressure of hydrogen outside*, and in addition there is the pressure of the nitrogen, so that the *total pressure in the tube is equal to the sum* of the original nitrogen pressure and that of hydrogen drawn in ; permanently greater, therefore, than the pressure outside. The semipermeability may in this case be attributed to the solubility of hydrogen in palladium ; the palladium septum, exposed



to hydrogen gas on one side, takes up some and dissolves it, just as, e.g., water exposed to ammonia gas takes it up: but, on the other side, there is no hydrogen to start with, so hydrogen evaporates from its solution in palladium, just as ammonia evaporates from aqueous solution on exposure to air, the process going on till the pressure of hydrogen on both sides is the same.

In the same way a moist animal membrane is semipermeable for certain liquids: if placed between ether and benzene the ether will get through into the benzene, because it is somewhat soluble in the water with which the membrane is saturated; but the benzene cannot get through into the ether, because to do so it would first have to dissolve in the water of the membrane; but benzene is insoluble in water.

The best-known semipermeable partitions for solutions are composed of precipitated ferrocyanide of zinc or copper, though the mode of action is not in this case clear. If a porous jar be filled with copper sulphate solution, and immersed in a vessel of potassium ferrocyanide, a deposit of copper ferrocyanide forms in the pores of the earthenware, and constitutes the required membrane, to which the earthenware adds mechanical strength. This partition allows water to pass through it pretty freely, but holds back salts and other dissolved substances, those of high molecular weight the most completely.

Suppose such a vessel filled with an aqueous solution, say of sugar, closed by a tight-fitting cork through which a long narrow vertical glass tube passes, and immersed in a beaker of water. Then the tendency of the water to flow through into the solution, which is not balanced by any tendency of the salt to diffuse out, causes the pressure inside to rise above that outside, just as in the case of nitrogen and hydrogen with a palladium septum. The difference of pressure, in fact, shows itself in a rise of the liquid up the glass tube till the hydrostatic pressure of the column of liquid in that is sufficient to prevent more water being drawn into the porous pot. Such an arrangement is called an osmotic cell, or osmometer, and the *excess of pressure in the solution over that in the water* is called the *osmotic pressure*. In the experiment on gases hydrogen is to be regarded as solvent (since it passes freely through), nitrogen the dissolved body; and the partial pressure of the nitrogen is the analogue of the osmotic pressure in the aqueous solution.

The phenomena of osmotic pressure can best be accounted for by means of the analogy that van 't Hoff drew between solutions and gases. In gases the pressure produced on the walls of a containing vessel is attributed to the movements of the molecules, which are practically free from each other's influence. Now in solutions, when dilute, the molecules of dissolved substances are so separated by the much greater quantity of solvent as to be mainly free from each other's influence, although of course subject to the influence of the molecules of solvent. For some purposes, however, the action of the solvent may be disregarded, and then the dissolved substance may be regarded as in a condition similar to that of a gas. For example, if a solution be enclosed in the osmometer, the dissolved substance tends, like a gas, to enlarge its volume, exercising a pressure on the boundaries containing it; it cannot escape from the liquid, not being volatile, but since water passes freely through the membrane of the osmometer, it can enlarge the space available for it by drawing water in from outside. This is what will happen in fact: water will pass into the osmotic cell, and the solution consequently extend higher up the tube, and increase in hydrostatic pressure. In this way the osmotic pressure of a solution may be measured, and it is found that the analogy between it and gas pressure is so close that Boyle's, Charles's, and Avogadro's laws are applicable to it. We may accordingly write

$$P = \frac{RT}{V},$$

where  $P$  = osmotic pressure,  $T$  = absolute temperature,  $V$  = 'dilution' or volume (c.c.) occupied by one molecule of dissolved substance (analogous to the molecular volume of a gas), and  $R$  is the same constant as in the gas equation (p. 79).

This equation is verified by experiment for organic substances such as sugar, but it is found that acids, bases, and salts give a greater osmotic pressure than is calculated in this way; so that we may put

$$P = \frac{R\iota t}{V},$$

where  $\iota$  is a factor which amounts, e. g., to about two in the case of potassium or sodium chloride. This excess of osmotic pressure is accounted for by a dissociation or break up of molecules of salts into two parts, the *ions* with positive and negative electric charges (see below, p. 188).

The cells of vegetable tissue, amoebae, bacteria, blood corpuscles and the like, behave like osmotic cells, the layer of protoplasm next the cell wall serving as a semipermeable membrane which allows water to pass through, but not salts and other dissolved materials. Hence, when a bacterium, e. g., is placed in water

the osmotic pressure within it draws in water till the cell wall is taut and the cell *turgid*. But if placed in a very strong aqueous solution, the osmotic pressure is greater outside than inside the cell, water is withdrawn from it, and the cell shrivels up, or becomes *plasmolysed*, and its vitality suspended, if not destroyed: thus condensed milk does not allow of the growth of bacilli in it, on account of the osmotic pressure of the sugar it contains.

It is to the osmotic pressure of the sap in vegetable cells that is attributed its power of rising to the top of high trees. The roots of a plant, embedded in damp earth, may be compared with the porous jar of an osmometer, the vertical tube of which is represented by the plant-vessels.

Certain apparently dissolved bodies, such as gelatinous silica and proteids, can be separated from water by means of a membrane of parchment paper or animal tissue. These bodies have received the name of *colloids*, and their separation from water is described as *dialysis*; to carry it out in practice, the liquid containing colloids and other substances is placed in a dish of parchment paper floating on water, and the water continually renewed by allowing a gentle stream from the tap to flow through it. After a time only the colloids remain. It has lately been shown possible to obtain many metals in a colloidal form in water; and it is probable that colloidal solutions are not really solutions at all, but merely suspensions of the solid in water, and therefore do not produce osmotic pressure.



## CHAPTER IV.

### PROPERTIES OF SOLIDS.

#### § 1. Elasticity of solids.

THE distinction between solids and fluids depends on their respective elastic properties, and may best be explained after some attention has been paid to the phenomena of elasticity in solids. The definitions given in reference to fluids (p. 80) may still be used, but with an extension of meaning. Thus in solids, a *strain* means a change either in volume or in shape; the latter class was not considered in dealing with fluids, because they have no tendency to recover from such strains, but will take freely the shape of the containing vessel. In solids, however, there is in general a tendency to recover from changes both in volume and in shape, and this tendency constitutes their elasticity. It must, however, be noted that the tendency is not unrestricted. Thus, if a piece of copper wire be slightly bent, on releasing it will spring back to its former shape; it is therefore *elastic*: but if it be strongly bent, it will remain bent after the force is removed; it has therefore reached an *inelastic* condition. *Stress* is, as before, the term used for the change in force producing a strain. With the aid of these conceptions, we shall consider first the behaviour of a typical, perfectly elastic, solid; and then the imperfections of elasticity that occur in solids when subject to larger strains.

(a) Strains in a solid, then, may be classified into (i) pure change of volume; (ii) pure change of shape; (iii) strains involving change of both size and shape. The first class follows the same laws as in fluids; the strains must be measured, not by the change in volume simply, but by the change occurring in 1 c.c. of the substance, or, in other words, the fractional change in volume: this rule is general, for the same reason as in fluids,

so that a strain is always a ratio of one volume to another, or one length to another. It is not practicable to produce a uniform expansion in a solid, but a uniform compression can be produced, as in a fluid, by applying a uniform (hydrostatic) pressure to it, everywhere perpendicularly to its surface: thus imagine a cube of material, exposed to a uniform pressure (say by enclosing it in a vessel and compressing the air in the vessel with a force pump), then it will become slightly smaller in each direction, but will remain a cube, so that the change is one in size, but not in shape. The change in size which it is possible to produce in a solid is always very small, and within these limits Hooke's law is true, viz. that the strain is proportional to the stress. The stress here, as in the case of a fluid, is the uniform pressure (reckoned per sq. cm. of area on which it is applied); and again, the constant ratio of stress to strain is called the *coefficient of volume elasticity*. The latter is therefore

$$= \frac{\text{stress}}{\text{strain}} = \frac{\text{pressure (dynes per sq. cm.)}}{\text{diminution in volume} \div \text{original volume}}.$$

The coefficient is even larger for most solids than for liquids;

e. g. for glass it is about  $4 \times 10^{11} \frac{\text{dynes}}{\text{sq. cm.}}$  or  $4 \times 10^5$  atmos. Hence

when an additional pressure (stress) of 1 atmosphere is applied to a piece of glass, it becomes smaller by  $\frac{1}{400000}$  part (strain). In this way the change in reading of deep-sea thermometers, due to the enormous pressure they have to stand, has been calculated. Organic tissues, though very soft, have about as much

resistance to compression as water; not so much as glass and metals it is true, but still so much as to make any effects of pressure in changing their volume usually quite negligible: it is only when a solid is permeated with air or other gas (like cork) that it is compressible to any considerable extent.

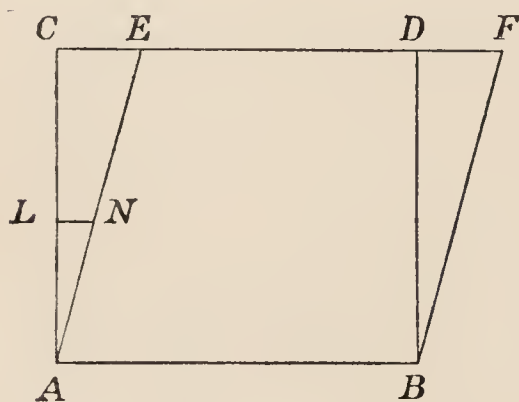


FIG. 44.

(b) Change of shape unaccompanied by change of size is effected by the kind of movement known as a *shear*: in this the material is displaced in parallel layers (Fig. 44), a rectangular piece of

it (ABCD) becoming distorted into an oblique parallelogram (ABEF); if the layer AB is kept fixed, the displacement of any point is proportional to its distance from the fixed layer; thus L is carried by the shear to N, and the shear is measured by the ratio of this displacement to the distance from the fixed plane, or  $\frac{LN}{LA}$ .

A familiar example of shearing is to take a book and slide one cover to the right and the opposite one to the left. To produce a shear, a shearing or tangential pressure is required (in the direction of CD and AB, Fig. 44), which again must be reckoned per unit area of the surface to which it is applied. Hooke's law is true here also, when the strain is small, so that the shear is proportional to the shearing stress applied, and the coefficient of elasticity in this case is called the *rigidity*, which is therefore

$$= \frac{\text{shearing stress (dynes per sq. cm.)}}{\text{shear}};$$

but whilst a solid that is compressed equally in all directions obviously cannot give way under the stress, this is not the case with shearing, and if too great a shearing stress be applied, the material will break. When a pair of steel shears or scissors is used, the force applied is distributed over a very small area, the area of the 'edge' of the shears, so that the shearing stress (per sq. cm.) is great, and the finer the edge the greater the stress becomes; consequently the material so stressed may be cut.

The difference between a solid and a liquid lies essentially in the possession of rigidity by the solid. If a shearing stress (within due limits) be applied to a solid, it merely causes it to be sheared to a definite extent, when the elastic reaction of the solid becomes sufficient to balance the applied stress, and equilibrium results. But if a shearing stress (however small) be applied to a liquid, since there is no elastic reaction, the liquid will give way indefinitely and a continuous flow take place. An example of this occurs when the wind blows over the sea; the friction of the wind on the surface of the water constitutes a shearing stress, and causes a flow or current to be set up in the water. Thus if Fig. 44 referred to a portion of liquid, AB being kept fixed by contact with a solid surface (the bottom of the sea), CD will not come into the position EF and stay there, but will



flow with uniform velocity in that direction so long as the stress is applied. The different layers of the liquid will in fact slide over one another in this direction; the motion is similar in character to that already described in the case of capillary tubes (p. 116), where the speed of flow is greatest in the centre and falls off gradually towards the outermost layer, which is held fixed by contact with the solid walls of the tube; and in both cases the motion depends for its amount on the viscosity of the liquid.

Returning to solids, we find extreme differences between the amount of rigidity they possess. Thus the coefficient for glass is about  $1.8 \times 10^{11} \frac{\text{dynes}}{\text{sq. cm.}}$ , which is nearly half as great as the volume elasticity of the same material: steel,  $8 \times 10^{11}$ . Substances like bone, whilst much less rigid than the metals, also have a coefficient of rigidity comparable with their volume elasticity: but pseudo-solids, like muscle and other soft tissues, jelly, india-rubber, have a degree of rigidity which is only a small fraction of their elasticity of volume; so that the small stresses occurring in the animal body are capable of producing large changes of shape in them, although the changes of volume due to equal stresses are, as was remarked above, negligible. Hence, when a soft tissue is distorted it may practically be considered as suffering a change in shape only.

There is another instance of strain which consists of change in volume only. When a cylinder is twisted, each concentric sheath of it is sheared. If we imagine a cylindrical sheath, slit down its length and unwrapped, it becomes a rectangle (Fig. 44, ABCD); if, however, it be first twisted, the slit comes to occupy an oblique position, and on unwrapping, the sheath becomes an oblique parallelogram (ABEF), that is, it is sheared. Consequently the resistance of a cylinder to twisting depends on the rigidity of the material of which it is composed.

(c) The most important kind of strain in practice is that produced by a longitudinal tension or pressure; an element of a structure subject to the former is called in engineering a *tie*, one subject to the latter a *strut*; thus a man's legs are struts, while, if he hangs on to a bar by the hands, his arms are ties.

The strain produced by such a tension or pressure is not a simple one, because both the shape and the volume are altered. Taking the case of a tie, for definiteness: if tension be applied to a metal bar, it elongates slightly, and still more exact measurement shows that it contracts very slightly in diameter; if the experiment be made on a cylinder of india-rubber such as a door-spring, the contraction in breadth can easily be seen. There is thus an obvious change in shape: and the joint effect of the increase in length and the decrease in breadth is to cause an increase in volume, so that a longitudinal extension really belongs to the class of mixed strains. However, the change in volume will be small when the rigidity of the material is small, as is the case with soft tissues, and so we come back to the fact mentioned previously, that in such tissues, though the shape may be much altered (as by stretching), there is no noticeable change in volume.

The extension (measured as a fraction of the original length) is in accordance with Hooke's law, proportional to the stress (tension) applied, again provided the strain be not too great; and the ratio of stress to strain in this case has received the name of Young's modulus, i. e.

$$\text{Young's modulus} = \frac{\text{stress}}{\text{strain}} = \frac{\text{tension per sq. cm.}}{\text{extension} \div \text{original length}}.$$

In the case of a strut it is only necessary to substitute the words pressure and compression for tension and extension: the value of Young's modulus is the same. A longitudinal compression is accompanied by a lateral expansion and a decrease in volume, which, however, in the case of soft tissues is again very small. The value of Young's modulus for a few substances is given in the table, p. 137.

From those numbers can be calculated the force required to produce a given extension, or the extension produced by a given force. Thus, suppose 100 gms. weight attached to a muscle of average cross section 0.4 sq. cm.; then the stress is  $\frac{100}{0.4} = 250$  gms. weight per sq. cm. The strain =  $\frac{\text{stress}}{\text{modulus}} = \frac{250 \text{ gms.}}{100 \text{ kilos.}} = 0.0025$ , i. e. the muscle will stretch by 0.0025 of its length.



When in place of a pressure from outside a tension from inside a muscle is applied (i. e. the muscle contracts), the effect is the same, viz. there is a decrease in length and an increase in thickness, while the volume remains very nearly the same.

When a beam is bent, whether by hanging a load to the middle of it, or when the beam is fixed into a wall, by hanging a load to the free end, the effect is mainly to stretch the side that becomes convex (the lower side of the beam weighted in the middle) and compress the other side; consequently the resistance to bending depends essentially on Young's modulus.

A fluid contained in any vessel exerts, we have seen, a pressure against all the sides of the vessel; this pressure is balanced by the resistance of the solid walls, and only when that balancing resistance is removed (as in a tap) results in motion of the fluid in flowing out. Now no solid is absolutely rigid and unyielding: even a steel vessel can only offer a resistance to an applied pressure, because it becomes slightly strained by the pressure; the resistance is therefore an elastic resistance, proportional (by Hooke's law) to the strain. If a liquid be contained in an ordinary solid vessel, the extent to which the solid yields is so small that only the most minute measurement can detect it. It is, however, occasionally necessary to take this yielding into account: thus in a mercury thermometer, held vertically, there is a considerable pressure on the inside of the bulb due to the column of mercury over it, which causes the bulb to expand: whereas if the thermometer be laid horizontally, this pressure does not exist; and in order to attain the highest possible accuracy in thermometry it is necessary to take this fact into account. But in soft solids (with low coefficients of elasticity) the effect is far more marked: thus an india-rubber bag becomes considerably distended if it be filled with water; and the same is true of arteries and veins. It has already been noticed that a fluid pressure can exist in an entirely closed vessel, having no free surfaces, such as the circulatory system: in this case the pressure on the fluid at any point is that exerted on it by the elastic stress of the distended vessels. The pressure of course increases downwards in the fluid, as in all cases; but even at the highest point in the vascular system there is not a free surface of liquid, but a pressure exerted by the walls of the



vessels, so that if a cannula were inserted there, the blood would rise to form a free surface considerably higher up.

In the same way the pressure in a gas, since the gas fills any vessel it is placed in, is balanced by the elastic reaction of the containing vessel, which is, therefore, strained by the pressure, more or less.

## § 2. Limits of elasticity.

Hitherto we have considered only the behaviour of solids so far as they are perfectly elastic, i. e. recover completely from a strain when the stress is removed. This is only the case, however, when the strain in a solid is kept within certain limits. What happens when those limits are passed is best illustrated by the consideration of a particular case, that of the longitudinal tension of a bar. When a steel bar is tested for engineering purposes, the effects observed are somewhat as follows. Let us suppose the bar to be a round one of 1.128 cm. diam. and consequently just 1 sq. cm. in cross section, and 10 cms. long. Then if a load of one ton ( $= 10^6$  gms.) be hung to the lower end, the stress in the bar is 1 ton per sq. cm.  $= 10^6$  gms. per sq. cm., or approximately  $10^9$  dynes per sq. cm. Substituting this number in the equation, Young's modulus  $= \frac{\text{stress}}{\text{strain}}$ , we see that the strain is  $10^9 \div (2.4 \times 10^{12}) = 0.00042$ .

This is the fraction of its original length by which the bar is stretched, but as it was originally 100 mm. long it becomes longer by  $100 \times 0.00042 = 0.042$  mm. This stretching can be measured by a very accurate pair of calipers. If the load be removed the bar will recover perfectly to its former length. If two tons be applied the stretching amounts to 0.084 mm., and again the recovery is complete, and the same is true up to about 2.5 tons. This behaviour is shown in Fig. 45 by the straight line  $OA$ ; so far

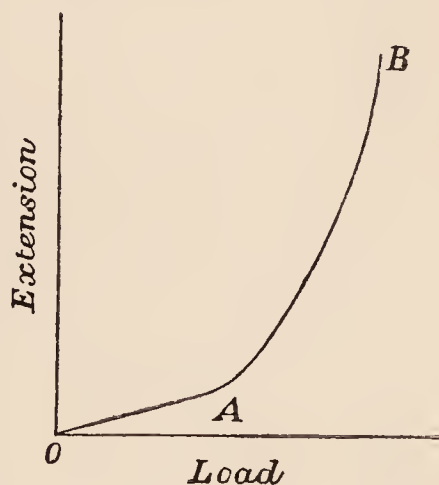


FIG. 45.

then the material is perfectly elastic, and follows Hooke's law (that the strain is proportional to the stress). But if loads

greater than about 2.5 tons are applied, the stretching is found to be out of proportion to the load, and then if the load be removed the recovery is only partial: the material has suffered *permanent strain or set*. The increased strain is shown by AB in the diagram. It will be noticed that this curve bends up rapidly: and whereas the stretching for  $2\frac{1}{2}$  tons is only about  $1\frac{1}{10}$  mm., that for 3 tons may amount to 2 or 3 mm., nearly all of which is permanent set.

The point A at which permanent set first appears is called the *limit of elasticity*. It is necessary in practice to keep a material within the limit of elasticity, since if used beyond that limit it becomes permanently altered in character and less able to support a further strain; if the process is repeated the permanent strain is increased, and so on till the structure breaks. Hence a load somewhat greater than the elastic limit, although not sufficient to cause fracture immediately, will do so if often enough applied and removed, and of course must be avoided in engineering structures; similarly a muscle or other organ may be injured by straining beyond its power of elastic recovery, although not actually broken.

When the stress is still further increased, the strain increases very rapidly, as the diagram shows, until eventually the breaking-point B is reached. This for an average specimen of cast-steel is at about  $4.5 \frac{\text{tons}}{\text{sq. cm.}}$ , or nearly double the greatest elastic stress. The corresponding strain may amount to 20 or 25 % of the original length.

The behaviour described is that typical of a tough and ductile material. A brittle one, such as cast-iron, glass, or bone, differs mainly in the absence of the phenomena represented by AB in the diagram: the limit of elasticity and breaking-point are practically identical, so that the material is elastic throughout, but fractures for a smaller load than a smaller tough substance, and does not stretch more than a very small fraction of its length before breaking.

There is another phenomenon shown by ductile substances, closely connected with permanent set, but not identical with it. Under extreme stresses the solid may come to flow like a liquid. When a steel bar is stressed beyond the limit of elasticity, it is only imperfectly elastic, it is true, but still the applied stress

is balanced by the reaction of the material, and equilibrium set up, so that the load may be left on for any length of time without further change. But when, e.g., lead is squeezed through a die, the amount of deformation produced depends essentially on the time during which the stress is applied: and goes on increasing indefinitely if the stress be kept up. This is the behaviour characteristic of a fluid, so that we may regard the solid as becoming, under great stress, a fluid of very great viscosity: it therefore flows very gradually (through the die), but when the stress is removed becomes solid again, and retains the shape which it has arrived at.

The breaking stress and Young's modulus for a few substances is given in the following table (in kilos. per sq. cm.):—

	Young's modulus.	Breaking stress.
Cast steel . . .	2,000,000	4,500
English oak . . .	120,000	400
Bone . . . . .	200,000	—
Nerve . . . . .	1,800	—
Muscle . . . . .	100	—

Hardness of a solid is a property distinct from its elastic behaviour, and is determined empirically by observing what substances it will scratch, or be scratched by; thus, diamond heads the list in the matter of hardness, emery and quartz are harder than glass, glass (usually) harder than steel. A hard material will wear down a softer one if rubbed on it; e.g. emery is used for grinding glass, and glass paper may be used for polishing steel.

Reverting to the general question of the distinction between solids and liquids, we see that only the former can remain in equilibrium under the action of a shearing stress; in a liquid under the same stress, however small, one layer will slide over another indefinitely so long as the stress is maintained. Hence, a solid has a tendency to retain its form, unless a very great stress is applied; but a liquid can be made to assume any shape under the action of any shearing stress, however small. There is, however, another difference which is of great importance. We have seen that when a solid is stressed beyond the limit of elasticity it suffers a permanent strain, and that if the stress is removed and reapplied it will be strained further, and



so on ; hence, the degree of strain depends not only upon the stress at the moment, but on the previous treatment of the specimen. This has been strikingly expressed by saying that the solid 'remembers.' Memory, in fact, in an organism consists in some enduring change, produced by previous impressions, which causes the organism to react differently to subsequent impression, and in this sense a steel bar may be said to remember, for its reaction to a tension (if beyond the elastic limit) applied for a second time is different to the first. This influence of previous history is attributed to a rearrangement of molecular groups caused by the stress, neighbouring molecules being brought into a strained position relatively to one another, and the influence may be removed—the memory washed out—by annealing, i. e. by maintaining for some time at a high temperature and very gradual cooling. At the high temperature, since the molecules are more mobile, it is easier for them to assume an unstrained position, and, if the cooling be conducted slowly, this position is not interfered with.

### § 3. Thermal properties of solids.

Solids, with one or two exceptions, expand on heating. The expansion may, as in the case of liquids, be expressed by means of a coefficient ; but whereas in liquids we are concerned only with changes in volume, in a solid we may have to do with changes in length too. Hence it is convenient to distinguish two coefficients ; thus, we may define the *coefficient of linear expansion* as the change in length per degree  $\div$  length at  $0^\circ$ , and the coefficient of volume expansion in the same way, with the word volume substituted for length. These quantities are always small, and it is of importance to notice what the relation between them is. To find this, imagine a cubical block of sides 1 cm. ; let it be heated till it measures 1.001 cm. in each direction, then its volume has become  $(1.001)^3$  c.c., which is very nearly 1.003. Hence, we may say that the expansion in volume is three times that in length.

The coefficients of expansion in length for ordinary solids vary from 0.000008 for glass and platinum to 0.000030 for zinc. Multiplying these numbers by 3 to obtain the coefficients in volume, we may compare them with what was said about liquids

on p. 90, and see that they are much smaller even than the corresponding value for mercury, which has the smallest expansion of ordinary liquids.

The effects of heat in solids, like those of stress, do not always correspond to the immediate cause, but are influenced by the previous history of the specimen of solid in question. Thus, there is a thermal after-effect well known in the case of thermometers. When a glass thermometer is raised to a high temperature, and afterwards tested in ice, it is found that the mercury stands rather lower at that point than it did before heating. This is because the glass bulb expands on heating and does not at once completely recover on cooling: the contraction may take days or even months to complete. This after-effect is especially noticeable in new thermometers, and is more marked with some kinds of glass than others.

An important consequence of the changes in size caused by temperature lies in the stresses that may be set up in solids thereby. If a solid bar be fixed tightly at both ends, and then heated, it will tend to expand, but, being prevented from doing so, will exert a pressure on the supports that are constraining it; and, similarly, a bar that cools and is not allowed to contract will be in a state of tension. The exact relation may easily be illustrated by an example: thus, suppose a steel girder to be firmly built into stone walls when at a temperature  $10^{\circ}$ , and subsequently to be heated by the sun to  $35^{\circ}$ ; then, if we take  $0.000012$  as the coefficient for steel,  $25^{\circ}$  rise of temperature will mean an increase of  $25 \times 0.000012 = 0.0003$  in its length. Young's modulus for steel is about 2,400 tons per sq. cm.; hence, if no allowance is made for expansion, and the ends be held quite fixed, a pressure will be set up in the girder equal to that required to shorten it by  $0.0003$ , i. e.  $0.0003 \times 2400 = 0.72$  tons per sq. cm.

When sufficiently heated most solids melt. The melting-point is perfectly definite in the case of substances of well-defined chemical composition, such as ice, naphthalene, sulphur, &c.; but bodies of an indefinite and mixed composition, like pitch or glass, turn gradually softer as the temperature rises, and cannot be said to pass from the solid to the liquid state at any one point on the scale. Usually fusion is accompanied by an increase in volume, so that when a substance is half melted the solid parts, being the denser, lie underneath the liquid; water is an exception to this rule, as it is to the corresponding rule of the expansion of liquids: ice, it is well known, floats on water, and this has important physiographical consequences, of the same kind as those discussed in connexion with the maximum density of water (p. 72).



A definite chemical compound melts and freezes quite constantly at a fixed temperature, so that whilst any solid is left unmelted, or any liquid left unfrozen, the temperature of the system will not change : this principle has been already referred to, for such a mixture of water and ice is used for standardizing thermometers in (p. 47).

There are some solids which on heating pass directly into vapour, a process known as sublimation ; this is met with in the case of camphor. Such cases do not, however, constitute a real exception, for they merely imply that the boiling-point, under atmospheric pressure, is lower than the melting-point. The boiling-point, we have seen, is raised by rise of pressure, and may in this way be brought above the melting-point : thus, if a piece of camphor be heated in a vessel containing air compressed to a few atmospheres, a little above the temperature under which it usually sublimates, it will melt instead, and then, a few degrees higher, boil.

These relations can best be understood by considering the vapour pressure of solids. That they have a vapour pressure is shown by many well-known facts, such as the slow evaporation of snow and ice during prolonged frost, and the smell of bodies like camphor. The pressure  $p$ , which is usually small, can be measured by the barometer method, as in the case of liquids, and may be represented by a curve drawn against tem-

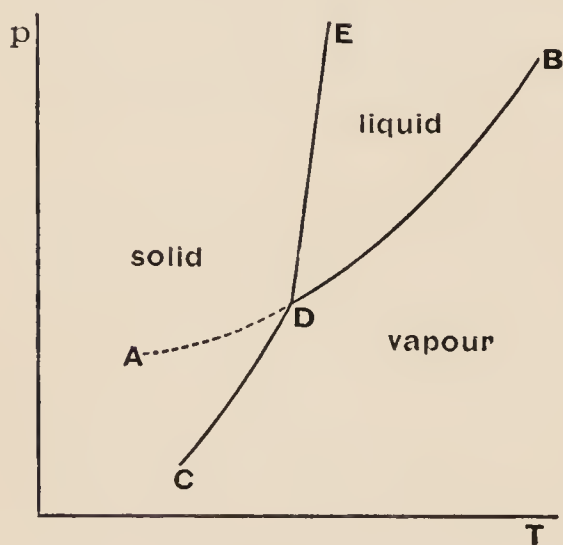


FIG. 46.

perature  $T$ ; the results obtained are such as are indicated in Fig. 46.  $CD$  is the vapour-pressure curve for the solid,  $DB$  for the liquid, and the point  $D$ , at which they intersect, the melting-point. Thus, for benzene it has been found that the point  $D$  corresponds to a temperature of about  $5^{\circ}.5$  and a pressure of 35 mm. If then solid benzene be gradually heated at atmospheric pressure it will melt at  $5^{\circ}.5$  nearly, and only on reaching a higher temperature ( $80^{\circ}$ ) attain a vapour pressure equal to that of the atmosphere and boil. But if the solid benzene be placed under the receiver

of an air pump, and be kept at 30 mm. pressure, on heating it will not reach its melting-point, but at about  $2^{\circ}.9$ , its vapour pressure having risen to that of the surrounding air, it will sublime. The apparently anomalous behaviour of camphor is therefore due to its pressure at the point  $D$  on the diagram being more than one atmosphere.



The diagram helps also to explain the phenomenon known as undercooling of liquids. If water be kept quite still and undisturbed in a vessel and cooled, it may be reduced considerably below  $0^{\circ}$  without turning solid. It is then said to be *undercooled*, and is not in stable equilibrium; for a small disturbance, and especially the addition of an ice crystal, however small, will cause a sudden freezing to occur, the temperature meanwhile rising to the normal level of the freezing-point: and the same phenomenon is observed in many other liquids. It appears, in fact, that a nucleus, round which crystallization can occur, is needed, in much the same way as in the condensation of vapour, the nucleus consisting often in dust from the air, or, most effectively, a particle of the liquid itself frozen. It has even been found possible to measure the vapour pressure of liquids in the undercooled condition, and the results are such as are indicated by AD (Fig. 46), i. e. the vapour-pressure curve for the liquid (ADB) is continuous above and below the freezing-point, while that for the solid (CD) corresponds to lower pressures than the liquid, and is not continuous with the liquid curve, but cuts it at an angle at D. Thus, solid benzene at  $0^{\circ}$  has a pressure of 24.61 mm., but undercooled liquid benzene 26.6 mm. This in itself shows that the liquid is unstable at that temperature, for if the two were placed side by side in flasks with a connecting tube the liquid would gradually distil over and condense in the solid form. Looked at in this light the melting-point is the *temperature at which the solid and liquid forms interchange stability*; only at the melting-point itself are they equally stable, and so can exist in equilibrium together.

The specific heat of a solid may be measured by the method of mixture, or by the steam, or the ice calorimeter, like that of a liquid; it is usually less than the specific heat of the same substance in the liquid state. The most important generalization in connexion with this subject is Dulong and Petit's law. This relates to the atomic heat, which is the specific heat of an elementary body multiplied by its atomic weight: this is found to be approximately 6.4 for most solid elements. Thus, copper with atomic weight 63 should have, according to this rule, specific heat  $6.4 \div 63 = 0.101$ : actually it is about 0.93. For further details we must refer to the books on chemistry.

When a block of ice is placed in a beaker, with a flame under it, it begins to melt, and for a long time its temperature remains stationary at  $0^{\circ}$ , although it is absorbing heat all the time. This heat is known as the *latent heat of fusion*. It is exceptionally large in the case of ice, amounting to 80 calories per gram; but all other solids on melting absorb more or less heat. This fact

may conveniently be illustrated by a diagram (Fig. 47) showing the relations between temperature and heat-content of 1 gram of substance. AB represents the gradual increase in the latter as the temperature of the still solid substance rises, the increased quantity per degree of temperature being the specific heat of the solid. At B the melting-point is reached, and here, without further rise of temperature, a large quantity—the latent heat—is absorbed, this being represented by the vertical line BC. At C the solid is all melted, and the rising line CD shows the effect of heat on the liquid, the slope of this line being more steep than that of AB, because the specific heat of the liquid is the greater, i. e.  $1^{\circ}$  corresponds to a greater absorption of heat.

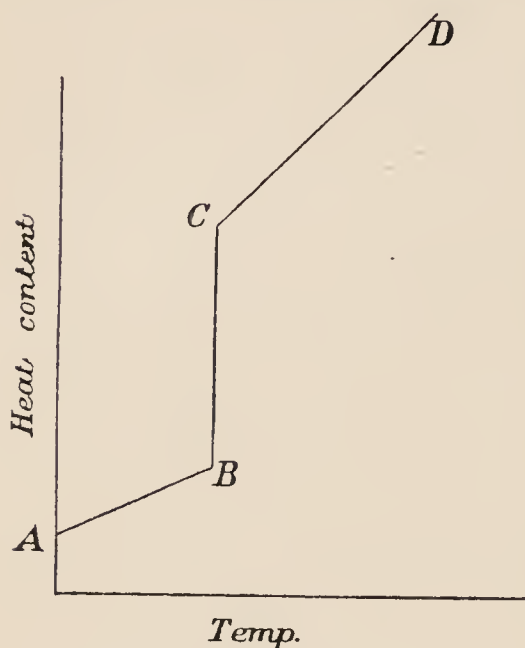


FIG. 47.

#### § 4. Solution.

When two liquids that mix in all proportions have freezing-points not very far apart, the behaviour of their mixtures with respect to solidification is normally that represented by Fig. 48. The figure may be taken to show the case of naphthalene and paratoluidine, which has been worked out. Naphthalene melts at  $79.3^{\circ}$ , paratoluidine at  $38.9^{\circ}$ ; but not only does a mixture containing a small amount of paratoluidine melt at a lower temperature than pure naphthalene, but the melting-point of the toluidine also is lowered by addition of naphthalene. The figure shows temperature vertically, and the percentage composition horizontally, so that the two points A and B stand for the two pure substances; AC indicates the melting-point of mixtures containing from 100 % to 31 % of naphthalene, BC mixtures of from 0 to 31 % of naphthalene, and it will be noticed that the freezing-point of each pure substance is lowered

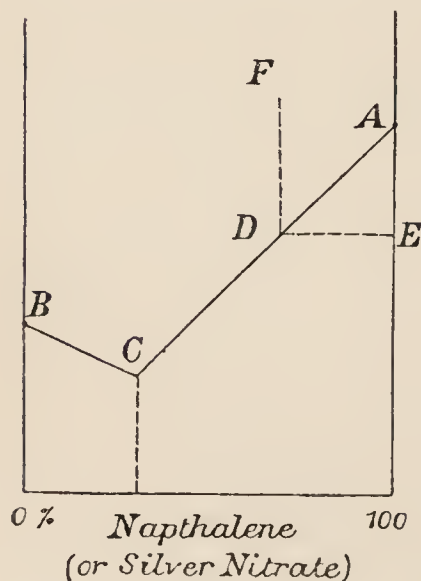


FIG. 48.

31 % of naphthalene, BC mixtures of from 0 to 31 % of naphthalene, and it will be noticed that the freezing-point of each pure substance is lowered



more and more by adding the other, till for a certain mixture (31 %) the point *c* is reached, which is common to the two curves, and the lowest freezing-point ( $29.1^{\circ}$ ) possible to the two substances: this is called the *cryohydric point*, and a mixture of the lowest possible freezing-point a cryohydrate (or sometimes a eutectic mixture). We may regard *ac* as belonging to solutions of paratoluidine in naphthalene, *bc* to solutions of naphthalene in paratoluidine.

The behaviour of this pair of substances is typical, not only of similar organic bodies, and of pairs of metals forming alloys, but also of the very important class of aqueous solutions. Consider, say, mixtures of silver nitrate and water. The melting-point of ice is  $0^{\circ}$ , of silver nitrate  $190^{\circ}$ ; at ordinary temperatures water dissolves a considerable amount of the salt, and as the temperature is raised the solubility is increased, but if we go on heating water in contact with the salt, in an open vessel, the experiment will be interfered with a little above  $100^{\circ}$  by the solution boiling. Let us then place the mixture in a closed vessel at high pressure, to prevent boiling; then it will be found that the water will take up more and more salt as the temperature rises, till at  $190^{\circ}$  the two substances, both liquid, will mix in any proportions. We have in fact been tracing out the line *ca* of the diagram, only instead of regarding it as representing the temperatures at which various mixtures freeze, we have looked at it in another aspect as showing the composition of a saturated mixture at various temperatures. Take, for instance, the point *d* on the line of saturation: this represents a certain temperature and a certain percentage composition, and we may reach it, either by starting with silver nitrate at that temperature and adding water to it till it is dissolved (i.e. proceeding along the line *ed*), or by making up a mixture of the required percentage at a higher temperature and cooling it till salt begins to crystallize out (along *fd*), the two processes leading to an identical result, viz. the solution, saturated with silver nitrate, *d*.

If, on the other hand, we have a solution containing very little salt, and it be cooled, then at a temperature somewhat below  $0^{\circ}$  it begins to freeze. But now it is saturated not with salt, but with the other constituent, and ice crystallizes out. This corresponds to the line *bc*, or to mixtures of a little naphthalene with much paratoluidine. Thus we see that the freezing-point of water is lowered by addition of salt, and that there must exist a cryohydric point (different for each kind of dissolved substance): common salt mixed in the right proportion with ice forms a cryohydrate melting at about  $-20^{\circ}$ , and is cooled to that temperature, forming a freezing mixture. The difference between the pair of organic solids quoted, and the pair water-silver nitrate, is chiefly in the fact that the melting-points of the latter lie so much further apart, and with other salts the difference is even greater.

The lowering of the freezing-point of a liquid, due to the presence of



another substance in solution, may be looked at in quite a different light. We have seen (p. 126) that the dissolved substance produces an osmotic pressure, which may show itself in drawing the solvent in through a semi-permeable partition, and would resist the converse process of removing the solvent in that manner. The osmotic pressure, in fact, shows itself in a tendency to resist removal of the dissolved substance in any manner, for by reducing the quantity of solvent one compresses the dissolved substance into a smaller space, and so increases the pressure which it exerts. It has already been noted that solution of a solid in a liquid raises its boiling-point (p. 99), i. e. adds a resistance to separation of the solvent in the form of vapour. Now, if a dilute solution—say of sugar in water—is cooled, the water begins to crystallize out, and the dissolved sugar is compelled to occupy a smaller volume: it follows, from the general principle referred to, that the freezing will only occur at a lower temperature than that of pure water. This *depression of the freezing-point* is, for dilute solutions, proportional to the osmotic pressure, and therefore to the concentration of the dissolved body (expressed in gram molecules per litre). It is, in fact, by measurements of the freezing-point that the osmotic pressure of solutions is mostly determined; and since the molecular weight (in grams) of any substance dissolved in a given volume of water produces the same lowering of the freezing-point, if we have a substance of unknown molecular weight, we may find this quantity by the effect produced on the freezing-point, e. g. molecular amounts of alcohol ( $C_2H_6O = 46$  grams) and urea ( $CO(NH_2)_2 = 60$  grams) produce the same lowering; if, then, we did not know the molecular weight of glycerine, but found that 92 grams of it produced the same effect as 46 of alcohol, we might conclude that the molecular weight of glycerine ( $C_3H_8O_3$ ) was 92. The same remarks apply to other solvents, except that the depression of the freezing-point due to a given concentration of dissolved substance varies from one solvent to another. But it must be borne in mind that the lowering is only proportional to the concentration for dilute solutions, because it is only to them that the laws of ideal gases are applicable (p. 127).

As was remarked in considering osmotic pressure, there is a large group of bodies which become electrolytically dissociated in aqueous solution, and so give anomalous results. Just as they give greater osmotic pressures than those calculated from their usual molecular formulae, so they give greater depressions of the freezing-point. Further explanation of this phenomenon must be postponed to the chapter on electricity.

## CHAPTER V.

### SOUND.

#### § 1. Sources of sound.

SOUND consists of waves propagated through the air, and received by the sensory apparatus of the ear ; it is, therefore, to be treated in accordance with the general principles of wave motion, considered in § 7 of Chap. I. As in that section simple oscillations were considered before waves, so here it will be convenient to treat first, for simplicity, the vibrations of bodies such as reeds and strings that generate sound-waves, making as few assumptions as possible as to the nature of the mechanism by which their motion is transmitted to the ear ; and leave for the subsequent sections those properties of sound which are more immediately connected with the behaviour of the medium through which it passes.

The vibrations to which sound-waves are due usually occur in solid bodies, and are nearly always due to elastic forces. In the preceding section we considered in some detail the motion of one vibrating elastic solid—a chronograph reed—chosen as a specimen of its class ; it may suffice here, therefore, to give a catalogue of the vibrating bodies that have most commonly to be dealt with : the motion of each is substantially the same in its harmonic or periodic character. Now, the most important kinds of strain of which a solid is capable are longitudinal strain (extension or compression), flexure, and torsion ; accordingly vibrating bodies may be classified as suffering these various strains. A bar or wire of metal when stretched tends to recover its former length, the tendency to do so being measured by one of the coefficients of elasticity, Young's modulus : the tendency to recover produces vibrations ; but on account of the great magnitude of Young's modulus for ordinary solids, these are usually too rapid to be used in music. Thus, a glass or steel



rod, clamped in the middle, or a steel wire, stretched between two pegs and stroked lengthways with a resined cloth, will emit a shrieking sound, the vibrations of which may be many thousand to the second. Such vibrations are longitudinal, i. e. each particle of the bar or wire vibrates in the direction of its length. But stretched wires or strings may also be made to vibrate transversely; this is the case which is of the most importance musically, including both piano and violin strings. A piano wire is struck by the hammer, and pushed sideways out of place, the ends remaining the while fixed on their pegs; consequently the wire becomes slightly longer, for the moment, and tends to contract again till it resumes its straight position. In this way vibrations are set up, each point of the string moving to and fro at right angles to the direction of its length. The frequency of these transverse vibrations is increased by making the wire more taut, so as to increase the force tending to restore it to its equilibrium position; the frequency is decreased, if the length or weight be increased so as to give it greater inertia.

The exact formula for frequency of vibration is  $n = \frac{1}{l} \sqrt{\frac{w}{m}}$  where  $n$  = number of vibrations per second,  $l$  = length (in cm.),  $w$  = tension of string (in dynes),  $m$  = mass of one cm. length of string (in grams). All transverse vibrations of strings, whether bowed as in the violin, or plucked as in the guitar, may be reckoned from this formula.

The flexural rigidity of a bar may be made use of to produce sound. The reed of Fig. 3 vibrates in that way, and really sets up sound-waves in air, but the vibrations are so slow as to produce no impression on the nerves of the human ear. If they be made more rapid by removing the weight  $w$  altogether, the spring will produce a faint, dull, booming sound, while a thicker bar would, on account of its greater stiffness, give more rapid and consequently more easily audible vibrations. A tuning-fork is essentially a pair of such flat bars, cast on the same base; its vibrations are more rapid the stiffer the prongs. They can be adjusted slightly in frequency by filing the prongs near the base (to make them slower) or near the point (to make them faster). The vibrations of flat and rounded plates (cymbals, gongs, bells) belong to the same class, though they are more complex in character.



Vibrations can also be set up by twisting rods and wires, but they are of no practical importance.

There is another class of musical sounds not fully accounted for in the description given above, viz. the sounds produced by wind instruments. In this class may be included the human voice and organ-pipes, as well as the orchestral instruments that are actuated by blowing. In all such cases the vibration is started by an elastic solid body—in the voice by the vocal cords, in organ-pipes by the reed, in the oboe or trumpet by the lips of the player; but the vibration is greatly intensified and modified in character by the air enclosed in the instrument. The way in which this air takes part in the motion may best be dealt with later, while considering the phenomenon of ‘resonance.’

(i) Musical sounds produced in any of these ways possess certain characteristics corresponding to those of the oscillations which generate them. These are, firstly, *loudness*: this is proportional to the total energy of the oscillation, and therefore depends on the amplitude. Now, by doubling the amplitude we double the force tending to restore an elastic body to its normal shape (Hooke’s law), as well as double the distance through which that force has to act; accordingly the work done by the force is quadrupled. We see, then, in general that the energy or loudness of a sound is proportional to the square of the amplitude of vibration of the sounding body.

(ii) Secondly, the *pitch* of the sound depends on the frequency of vibration. Notes of high pitch are those due to rapid oscillations, notes of low pitch to slow ones; while a noise, which is not a musical sound at all, is a sound corresponding to oscillations of so irregular a character that it is not possible to assign any definite pitch to it. Pitch is reckoned by intervals from some fundamental note—a process of a purely musical character—and it is found that a constant interval in the musical sense corresponds to a fixed ratio, not a fixed difference, in frequency of oscillation. Thus, the interval known as an octave corresponds to the ratio 2:1. If a certain low note makes 80 vibrations per second, the note an octave above it makes 160; if a certain high note makes 1,000 vibrations per second, the note an octave above that makes 2,000, although the difference in frequency is so much greater in this case.

The intervals of most importance in music are those expressible by the simplest numerical relations, viz. the octave 2 : 1, the fifth 3 : 2, and the major third 5 : 4. Out of these intervals, taken in various combinations, is made up the scale or succession of notes required in connexion with a starting or key-note arbitrarily chosen. The following table gives the diatonic major scale, which may be regarded as the most natural one for singing when not accompanied by any instrument whose mechanical construction gives a bias in favour of any more artificial arrangement. The key-note chosen is C, and the vibration numbers in the last column are those that would be obtained by starting from the middle note of a modern piano.

Note.	Interval.	Ratio of frequencies.	Actual frequency (Just intonation).
C	Key-note	1 : 1	264
D	Second	9 : 8	297
E	Major third	5 : 4	330
F	Fourth	4 : 3	352
G	Fifth	3 : 2	396
A	Major sixth	5 : 3	440
B	Major seventh	15 : 8	495
C	Octave	2 : 1	528

Modern music requires that the key-note may be changed to any point of the scale originally adopted (modulation), and hence a larger number of notes within the range of an octave is indispensable. A compromise between all the notes that might possibly be needed, consisting of twelve notes, has been adopted in practice; and in modern instruments with fixed pitches (piano, organ) the arrangement known as *equal temperament* is aimed at. This consists in making the interval (i. e. ratio of frequencies) between each two successive notes the same. Hence, each interval—called an equal semitone—is represented by the ratio  $2^{\frac{1}{12}} : 1$ , and, starting from C as above, the frequencies are 264,  $264 \times 2^{\frac{1}{12}} = 279.7$ ,  $264 \times 2^{\frac{2}{12}} = 296.3$ , and so on. In contradistinction, the notes of the natural scale given above are said to be in *just intonation*.

(iii) Sounds of identical pitch differ also in *quality*, as for instance the same note played on a piano, a violin, and a horn. This is a distinction which cannot be accounted for by any of the properties of simple harmonic motion, and indeed it is found experimentally that most musical sounds, while harmonic, are not simple. A tuning-fork gives vibrations of practically simple harmonic character, and accordingly when a tracing of them is obtained by means of the chronograph it is found to be a curve of sines (Fig. 8). If some other source of sound, such as



a violin string, be made to record its movements in an analogous way, the curve obtained will be harmonic, i. e. will repeat itself at regular intervals, but will not be of the same shape as for a tuning-fork. Innumerable different shapes can be obtained, some of them differing as much from a curve of sines as the sphygmograph curves already referred to (p. 35). These differences in harmonic character may be explained by saying that the vibrations of a body like a violin string are compounded of several S. H. M.'s arranged in a particular sequence: thus the string may vibrate as a whole—this gives the so-called *fundamental tone*, by which its pitch is determined. But the string might divide in two at the middle point, and the two halves vibrate in opposite phases; this can actually be done by holding the finger very lightly on the middle of the string while it is being bowed. Now each half of the string will emit separately a tone of twice the frequency of the fundamental (for the frequency of a string is inversely as its length, p. 146), i. e. a tone one octave higher; this is called the second *partial tone or harmonic*. Again, the string will break up into three, if the finger be put lightly at one-third of the length; the vibrations of each segment have then three times the frequency of the whole string, and their pitch is an octave and a fifth above the fundamental tone; this constitutes the third harmonic, and so on. Now when a violin string is bowed in the ordinary way, its motion is not that due to the fundamental S. H. M. alone, but the various other partial tones occur at the same time, and the actual course of each point of the string is a complex harmonic motion.

The analysis of a complex periodic motion into a succession of S. H. M.'s with frequencies in the ratio 1 2 3, &c., was first given by Fourier, and the series is often known by his name.

We can now see that the different relative strengths of these partials will produce differences in quality between two sounds, although the fundamental tones of the two may be the same in pitch. Actually the flute pipes of an organ possess only the second, and faintly, the third harmonics; violin strings the first eight or ten, in regularly decreasing amplitude; piano strings only the first six, but the second, and even the third, sometimes more strongly than the first; while in the note of a trumpet even the twelfth or fourteenth may be detected, and so on.



The more important musical instruments, from their construction, are capable of producing partial tones whose frequency, as just shown, is an exact multiple of that of the fundamental tone : these are called *harmonics*. But vibrating bodies of more complex shape, such as bells, possess, it is true, more than one way of vibrating, but the frequencies corresponding to the various modes do not bear this simple relation to one another. The partial tones, i.e. the simple harmonic motions into which the actual vibration can be analysed, are said in this case to be *anharmonic*.

## § 2. Sound-waves.

A vibrating body, such as those described in the preceding section, acts as a source of waves in the medium—usually air—around it. Thus, to take a tuning-fork as an example, when its prongs move outwards it pushes against the surrounding air, tending to compress it; the air cannot move outwards instantaneously, although it does so very rapidly, consequently the air in the immediate neighbourhood of the fork is actually compressed for the moment. But air has elasticity of volume, i.e. tends to recover its size after compression; so, as the compressed air is at a slightly higher pressure than the undisturbed air outside, it will tend to recover its normal volume by compressing in turn the next layer of air. Thus the compression is passed onwards from each layer to the next. But very soon the prongs of the fork are moving in again, leaving room for the air in immediate contact with them to recover its volume, and even expand. This change is in a similar way communicated outwards to one layer of the air after another; and so it comes about that the complete oscillatory motion of the air—condensation and rarefaction alternately—is propagated outwards from the fork, and each part of the air suffers in turn changes similar to those of the air first acted upon : this constitutes the wave.

Waves may be either *longitudinal* or *transverse*, i.e. the direction in which the vibrating particles move to and fro may be lengthways or crossways to the direction in which the wave is proceeding through the medium; but in any particular case it is only possible to generate the type of wave for which there is the necessary elasticity in the medium. Thus, in the experiment with a rope (p. 35) the motion of each point is obviously transverse to the direction of the rope itself, and the latter constitutes the medium; this is possible because of the degree

of rigid connexion that exists between successive portions of the rope; although flexible, a shear in any section of it tends to cause a shear in the next section, otherwise the rope would go to pieces. If instead of a rope we had a cylinder of sand, this rigidity would not exist, and if one section of the sand were displaced sideways, it would have no tendency to drag the next section after it.

Now, air has no rigidity, but it has elasticity of volume; consequently transverse waves are not possible in it, and the waves of sound are longitudinal. These longitudinal waves may perhaps be more easily imagined, if their propagation along a straight tube of air be considered; it is then obvious that rarefaction and condensation of the air in the tube will be produced as the air moves to and fro.

When waves are propagated in one direction only (as along a tube), they do not diminish in amplitude (and therefore in intensity) except so far as absorbed by frictional resistance. But when propagated in space, as the wave spreads outwards the energy it carries is spread out over a larger and larger spherical surface, in proportion to the square of the distance from the source, and consequently the energy at any one point (or intensity of the sound) is inversely proportioned to the square of the distance from the origin.

In a wave the frequency of vibration at any point is necessarily the same as that of the source which is producing the wave; but there is another characteristic quantity—the wave velocity—which depends on the nature of the medium through which the sound is passing, and, in the case of sound-waves at any rate, depends on that only.

We may form a clearer notion of the velocity of a wave and the associated quantity, the wave-length, by considering in more detail the mode of propagation. The five curves (Fig. 49) are intended to show this. No. 1 may be taken as representing the condition at a certain moment of a portion of air (say in a tube) through which a sound-wave is passing. The air at the point A is at that moment at its normal (equilibrium) position, but moving towards the right; that of the point B, less distant from the source of the wave, is in a subsequent phase of its motion, viz. as far towards the right as it can go; for the condition that B is in now is that towards which a point further on such as A is tending. Again, C is in a more advanced phase still, that of passing through the centre



towards the left ; *D* is at the extreme left ; while *E* is in the same condition as *A*, the cycle of phases being complete. Next consider No. 2 ; this is drawn for the same points, but at a moment later by one quarter period. Consequently the air at *A*, which was moving through the centre towards the right, has now reached the extreme right—the phase that *B* was at in the earlier diagram ; similarly, *B* is in the phase formerly held by *C*, and so on. So that the total effect is as if the arrangement No. 1 were shifted forwards through the distance *BA*, but otherwise unchanged. Similarly, Nos. 3, 4, 5 represent the same wave at moments later than No. 1 by  $\frac{1}{2}$ ,  $\frac{3}{4}$ , and 1 period, and each one shows the wave advanced by the same amount, as compared with the preceding figure. We have, then, a picture of the propagation of the wave, and see that the distance between two points in the same phase, such as *A* and *E*—the *wave-length*—is the distance that the wave travels in the interval between Nos. 1 and 5, that is in one period.

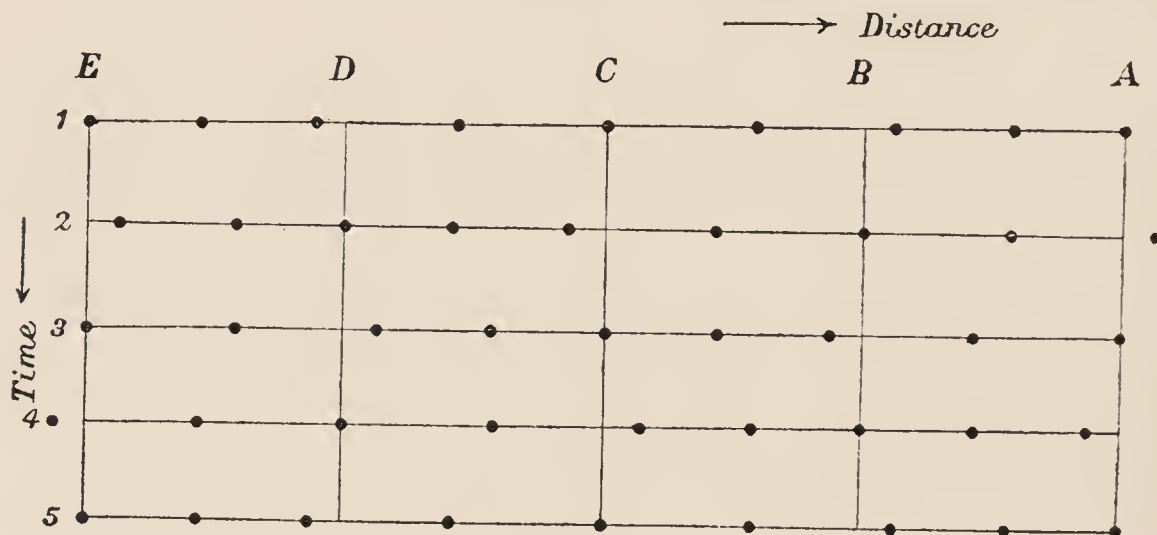


FIG. 49.

Hence the relation already given (p. 36) that the wave-length = wave velocity  $\times$  periodic time.

We have so far described the sound-wave in air by means of the *displacements* of the air particles which it produces. It may also be looked upon as a wave of compression and expansion, i. e. of change in *density*; but it is to be noted that the phases are different according to these two ways of looking at the phenomenon. Thus, looking at No. 1 in the foregoing figure, it appears that at *A*, *C*, and *E* there is no displacement of the air ; at *B* and *D* the maximum amount of displacement. But the air particles are crowding in on *A* and *E* from both sides, spreading out from *C* on both sides, consequently *A* and *E* are points of maximum compression, *C* of maximum expansion, while it is at *B* and *D*, the points where the greatest change of position occurs, that the density possesses its normal value.

Fig. 49, it will be seen, if read horizontally, indicates the state of all the air at an instant, i. e. the nature of the wave at that instant ; but if



read vertically it shows the changes of position or oscillation of a single air particle in time.

The velocity with which a wave is propagated in air or other medium depends on the density and elasticity of the medium. It is easy to see in a general way that increased elasticity would cause each layer of air to recover more quickly from the compression it suffers, and consequently transmit the motion more rapidly to the adjacent layer; while increased density would mean a greater mass to move, and consequently the elastic forces would be slower in producing their effect, and the rate of propagation would be less. Exact treatment of the problem shows that the

$$\text{velocity of sound} = \sqrt{\frac{\text{coefficient of elasticity of the medium}}{\text{density of medium}}}.$$

Now, it has been shown (p. 81) that the coefficient of elasticity of a gas, under the condition of constant temperature, is equal to its pressure. Applying this result to air at normal temperature and pressure, we have

$$\text{coeff. of elast.} = 1,013,200 \text{ dynes per sq. cm. density} = 0.001293 \text{ gms. per c. c., whence the velocity of sound} = \sqrt{1013200 \div 0.001293} = 28000 \frac{\text{cms.}}{\text{sec.}}$$

But the velocity has been determined by direct experiment, the best results being those of Regnault, who observed the time taken for the sound of a pistol shot to travel through long drain-pipes in Paris. His

$$\text{result is } 33060 \frac{\text{cms.}}{\text{sec.}}. \text{ The large discrepancy between the calculated and}$$

observed numbers was first explained by Laplace (before Regnault's time) as follows: the elasticity used in the above calculation is that derived from Boyle's law, i. e. on the assumption of constant temperature. But in a sound-wave the compressed portions are momentarily raised in temperature by the compression, the expanded portions cooled; and the movement is so rapid that there is no time for equalization of temperature by conduction. Hence, the elasticity ought to be calculated on quite another assumption, viz. that of no interchange of heat (the so-called adiabatic condition): now the adiabatic elasticity of air is some 40% greater than the isothermal elasticity; this accounts for the greater velocity of sound satisfactorily.

The velocity of sound is greater in hydrogen than in air, less in heavy gases; the velocity in water is some 150000  $\frac{\text{cms.}}{\text{sec.}}$ , and in some solids equally high.

From the velocity of sound the wave-length may be calculated;

e. g. the middle C of the piano, with a frequency 264 per second, has wave-length  $33060 \div 264 = 125$  cms., while higher notes have shorter, lower notes longer wave-lengths.

When sound-waves strike against an obstacle, they are *reflected* and travel backwards with the same velocity as before. If they meet the obstacle obliquely, we may consider the velocity of the wave resolved into two parts, one perpendicular to the obstacle, the other parallel to it; the former is reversed by

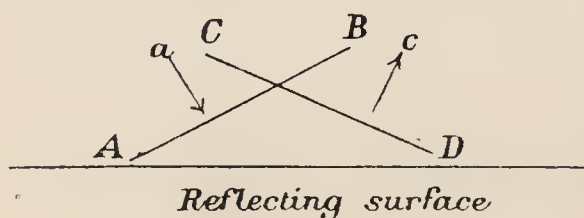


FIG. 50.

the collision, the latter goes on unchanged. The consequence of this is shown in Fig. 50. The incident wave-front AB travelling in the direction of the arrow *a*, is in this manner converted into the reflected wave-front CD travelling in the

direction of the arrow *c*, equally inclined to the reflecting surface, but in the opposite sense. Reflection of sound produces the well-known phenomenon of echo.

When sound passes from air into another medium in which its velocity is greater or less, it is in general bent out of its course; in the former case the wave-front comes to make

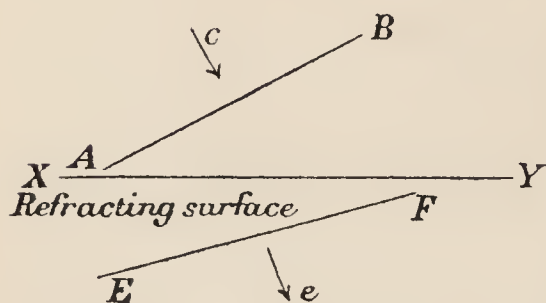


FIG. 51.

a greater angle with the surface bounding the two media, in the latter case a less angle (Fig. 51); this process is known as *refraction*, and the bounding surface XY in the figure is called the refracting surface. The angle between this and the incident wave AB is called the angle of incidence, that between XY and the refracted wave EF

the angle of refraction, the arrows *c* and *e* indicating the directions in which these waves are travelling. These two angles are only equal when they are both zero, i. e. the wave is parallel to the refracting surface; it then proceeds without change of direction.

The reflection and refraction of sound possess no great interest in themselves, but they are identical in character with the corresponding phenomena shown by light, and which are of great importance. The experiments of R. T. Wood illustrate



the behaviour of sound-waves very clearly, and his figures are reproduced here, not so much for their own sake as for reference when dealing with refraction and reflection of light. The method of experimentation was to make a loud electric spark ; this sets up a single wave in the neighbouring air, or what may be called a sound-pulse (see p. 35), like a stone dropped in a pond. Now, the condensation produced by a sound-wave would make the air faintly visible, in just the same way as the streaks of hot air over a flame are, were not the movement too quick for the eye. Accordingly another electric spark is produced, a very small

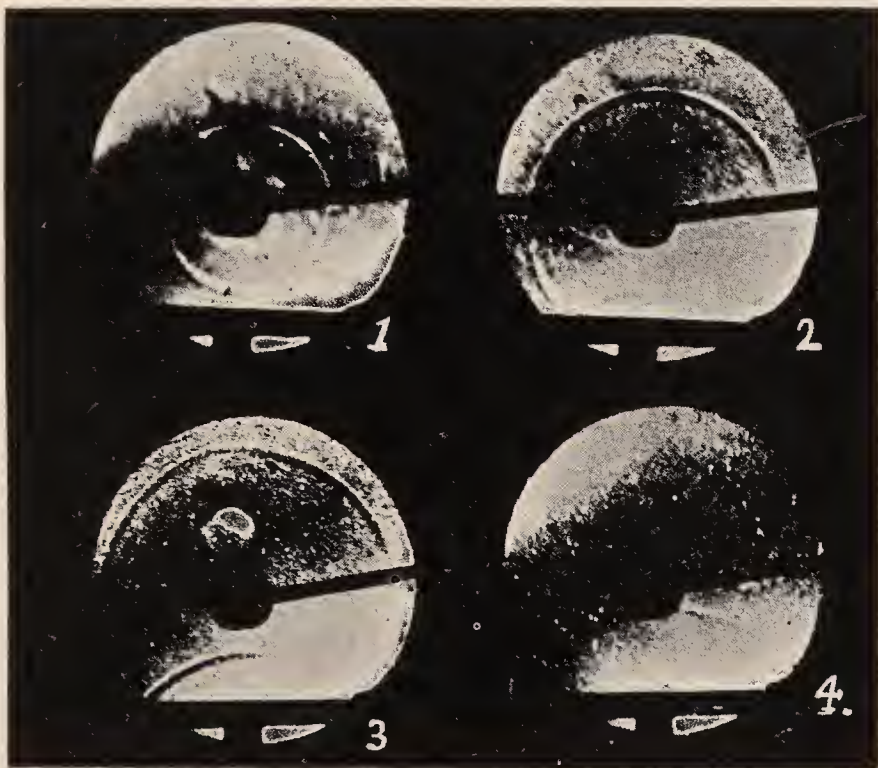


FIG. 52.

fraction of a second after the first, and by means of this the sound-wave is photographed in the position it then occupies. The waves could be caused to suffer reflection and refraction in various ways before being photographed. In the photographs (Figs. 52, 53, 54) the round black patch with a pair of arms to it represents the two knobs between which the spark took place ; from this the wave will be seen spreading out spherically till (Fig. 52, No. 2) it strikes a flat reflecting surface near the bottom of the figure ; by this it is inverted and the reflected portion is seen (Nos. 3 and 4) travelling back towards the balls. In Fig. 53



the reflection takes place at a spherical surface, causing the wave to bend sharply inwards (No. 2) and approach a point (3, 4).

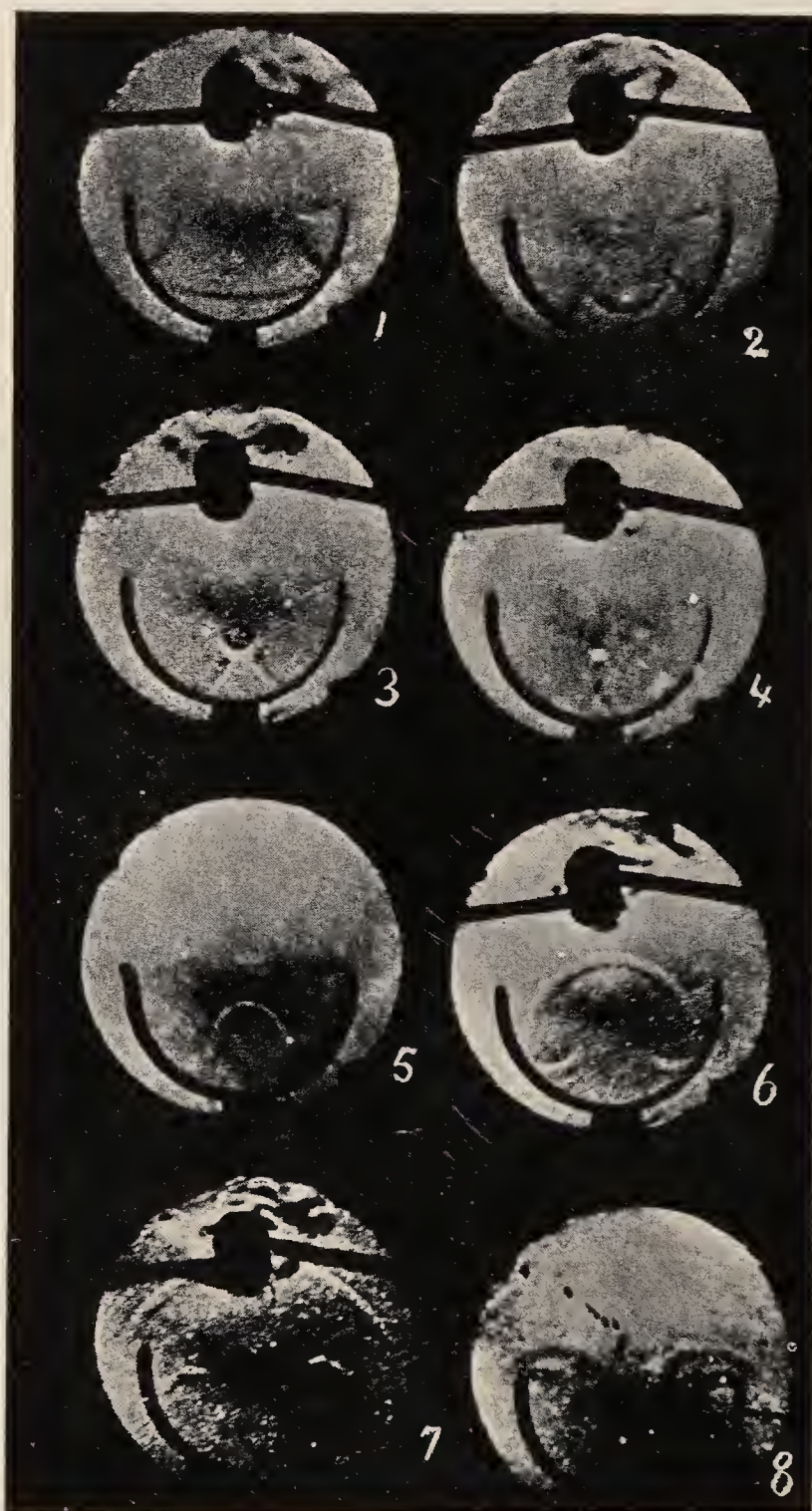


FIG. 53.

This point, on which the whole disturbance is for the moment concentrated, constitutes what, in optics, is known as the focus; after passing the focus the wave spreads out again spherically

(5, 6, 7, 8). The end portions of the reflected wave, which are prominent in Nos. 6, 7, 8, are introduced to illustrate the phenomenon of spherical aberration (vid. inf. p. 279). In Fig. 54 the lower part of each photograph is occupied by a vessel containing carbon dioxide (enclosed by a collodion film), in which sound travels less rapidly than in air. Here part of the dis-

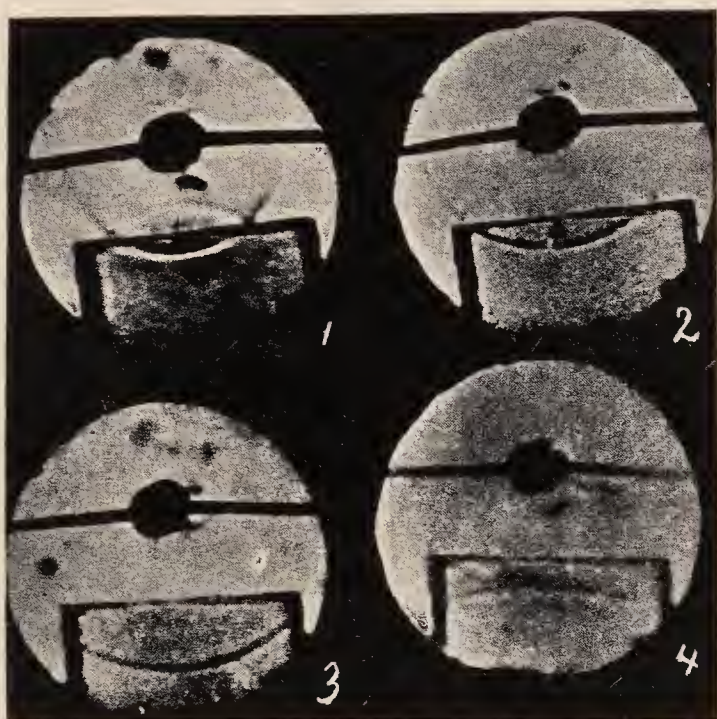


FIG. 54.

turbance is reflected, but the greater portion travels onward and suffers refraction; this is clearly seen in Nos. 1, 2, where the refracted portion is seen to have travelled a less distance than the reflected. In No. 4 the wave has struck the bottom of the vessel of carbon dioxide and is travelling upwards again.

### § 3. Resonance.

When a note is sounded, and there exists in the neighbourhood any other body capable of emitting the same sound, that body will be set in vibration too. This phenomenon of sympathetic vibration is known as resonance. It is by no means restricted to the domain of sound, but plays a very important part in connexion with vibrations of all kinds, in mechanics, in light and electricity. We may perhaps best illustrate the phenomenon,



in the first instance, by a simple mechanical example. Suspend a common pendulum, and give to the bob a succession of blows, each one so slight as to produce only a very small movement by itself. If these blows arrive at irregular times they will in general produce very little effect; if they are made at regular intervals, too, they will usually not set the pendulum swinging far, for on the average half of them will be made when the pendulum is moving in the direction opposite to that in which the blow would drive it, i. e. half the blows will help, half will hinder the vibration. But if the blows are so timed that, after the first one has produced a complete swing to and fro, the second is given, it will aid the swing set up by the first; and then the third, given one complete period later, will again increase the effect, and so on; i. e. the impulses must be periodic and have the same periodic time as the pendulum. In this case, even small impulses will in time set up a large vibration of even a heavy pendulum. The pendulum may then be said to be in resonance with the applied impulses.

To take, then, an instance from sound, let the damper of a note on the piano be raised, and the note sung loudly and accurately in tune; on suddenly stopping the sound of the voice the note will be heard to continue—the piano-string has taken it up, and vibrates for some time after the exciting cause has ceased. In this case, then, the voice produces sound-waves in the air, and these fall on the piano-string: each vibration of the air in contact with the string gives a small impulse to it, but, on account of the lightness of air, only a very small one. If these impulses were not properly timed, so that some helped and others hindered the motion of the string, the effect produced would be negligibly small, i. e. if the note be sung a little out of tune the experiment will fail; but if they are accurately in unison with the natural time of vibration of the string, their effects will all be in the same direction, and the amplitude of vibration will steadily increase until the string loses as much energy per second by giving out sound on its own account, as it gains by absorbing the sound-waves from the voice. If the resonant body be one that loses its vibrations slowly, i. e. suffers little damping, such as a tuning-fork not mounted on a box, the applied impulses must be very accurately in tune with it to produce resonance, but then the effect produced is



large; e. g. if two tuning-forks be of precisely the same frequency, and be pressed upon the same table, one of them on being struck will set the other vibrating quite loudly. But if the resonant body be subject to much damping, as the skin of a drum, the vibrations of which die out very quickly, the resonance becomes both weaker and more indefinite, i. e. the exciting cause may be appreciably out of tune with it and still produce some resonance; but even if accurately in tune it will not produce much. Moreover, bodies like a drum-skin, or still more those of irregular shape, such as the body of a violin or the sounding-board of a piano, possess many different modes of vibration, each with its natural period; so that, what with this and the indefiniteness of their resonance, they are capable of taking up to some extent any sound-waves that act on them. In the case of the piano or violin the vibrations are communicated directly from the string to the sounding-board, and the latter gives them out to the air more powerfully than the string alone would, on account of its much greater breadth of contact with the air. On the other hand, the drum-skin of the ear offers an important instance of the absorption of sound-waves from the air, its purpose being to transmit the vibrations to the interior mechanism of the ear.

Air-chambers constitute one of the most important classes of resonators. As an example of a kind easily dealt with, we will take an open tube, such as an organ-pipe. It is only necessary to blow over the mouth of such a tube in order to produce from it a faint sound of definite pitch. If the pitch be measured and the corresponding wave-length in air be calculated, it will be found to be double the length of the tube. This fact helps one to understand what is taking place in the tube. Thus, let us suppose that instead of merely blowing over its mouth, we hold there a vibrating tuning-fork of the same pitch as the natural vibrations of the pipe; it will then be found that the air of the pipe will be set in vibration much more strongly than before. Now consider the impulse given at a certain instant by the prong of the fork moving towards the mouth of the tube; this tends to compress the air in front of it, and sets up a wave which travels down the length of the tube. At the far end the wave meets the great mass of undisturbed air outside, and although some of its energy travels on and

becomes an ordinary sound-wave in the free air, some of it is reflected on account of the inertia of the air, and travels back along the tube to its mouth. By the time it reaches the mouth again it has travelled through twice the length of the tube, i. e. through one wave-length, and just one complete period of the fork's vibration has elapsed. Consequently the fork is again in the original phase and ready to impart another small impulse to the air, which helps the previous movement; so that in the course of a few vibrations the air in the pipe is set in vigorous movement. It appears, then, that a long thin pipe, open at both ends, will resound to a note whose wave-length is twice the length of the pipe. If the pipe be closed at one end, it acts like half of an open pipe, and the wave-length is four times its own length.

An organ-pipe is 'blown,' or set vibrating, by a small current of air entering at one end. This current does not traverse the length of the pipe, but merely serves to set in motion a reed fixed in the end (reed pipe), or if there be none, the 'lip' or thin lamina of wood or metal near the front aperture of the pipe. The air in the pipe is then set in sympathetic resonance with this reed or lip, and the tone produced is greatly strengthened.

Air cavities of other shapes are also capable of resonance. Their natural pitch cannot be calculated so easily, but it may in each case be found experimentally by blowing over the mouth. Thus, tuning-forks are sometimes mounted on open wooden boxes of appropriate size, in order to strengthen the tone which they give out. In the same way the cavities of the chest and mouth act as resonators to the sounds produced by the vocal cords.

The action of the resonance chambers in the latter case is of especial interest. We have seen that the distinctive character of notes from different musical instruments is due to the varying intensity of the upper partial tones they possess. This is true also of the vowel-sounds given out by the human voice; but whereas in a violin or horn the partials form a harmonic series with a definite relation to the fundamental note, in the voice that is not the case; but each upper partial is approximately fixed in pitch, being that proper to one of the air chambers in the chest or mouth. These are, no doubt, modified a little in accordance with the fundamental note produced by the voice,



so as to bear a harmonic relation to it ; but the general character of the vowel-sounds is that of a fundamental tone, variable in pitch according to the range of the voice producing it, modified by upper partial tones of approximately constant pitch. Thus, the sound *u* (Italian) or *oo* (English) is due to a nearly simple harmonic motion, but reinforced by the note F in the bass clef (176 vibrations per sec.) ; if the mouth be put into the shape necessary for speaking this vowel and a tuning-fork of 176 vibrations be struck and held in front of it, resonance will be produced ; for other vowels the mouth has to be altered in shape, and the resonance is modified accordingly.

With the aid of resonators any desired tone may be clearly picked out from a complex of sounds. For this purpose the most convenient form is a glass globe, with a small opening at one side to insert in the ear, and a larger mouth on the opposite side to turn towards the source of sound. Any desired pitch may be obtained by choosing the appropriate size. If such a resonator be held to the ear, the corresponding note can always be detected in a confused noise, such as the sound of traffic in the streets ; whilst if a note on the piano be sounded—say C with 132 vibrations—and a resonator of 264 or 396, &c., be held to the ear, the second, third, &c., harmonic will be easily distinguished from the fundamental tone.

#### § 4. Perception of sound.

The physiology of hearing is still in a rudimentary state : even the general principles on which the auditory mechanism acts can scarcely be said to be settled ; it would therefore be inappropriate here to discuss at any length the theories that have been proposed. We shall merely mention certain facts in connexion with the subject, and such deductions as are most free from doubt.

To begin with the range of hearing, it should be noted that vibrations, which are incapable of appreciation by the ear, may be of the same character as those which produce the sensation of sound. It is, in fact, only within certain limits of pitch that waves in air are perceptible to the ear ; many observations have been made on the upper and lower limits, giving somewhat discrepant results, since the sensitiveness of different ears is not the



same ; but, as an average, the range may be put at from 33 vibrations per second (the bottom C of the piano) to about 16,895 vibrations (two octaves above the top note of modern concert pianos), or a total range of nine octaves ; notes lower than about 33 vibrations, on the piano and organ, are heard almost entirely by means of their upper partials. In the central part of this range of sounds the ear is capable of distinguishing, with practice, differences of pitch as small as one-twentieth of a semitone ; at the ends of the range the sensitiveness is less, but in any case more than a thousand grades of pitch, between the highest and lowest, can be distinctly recognized.

Vibrations can be produced which are much slower than 30, or much faster than 16,000 vibrations to the second, and which, being quite similar in physical character to the vibrations of sounding bodies, give out waves to the air in all respects except pitch identical with sound-waves ; the distinction implied by the term sound is therefore a purely physiological one. Very slow vibrations, such as those of a chronograph reed (Fig. 3), are perceived as a throbbing, the successive waves not blending in their effect on the nerves ; when the speed is increased this throbbing sensation gradually passes over into a dull booming sound.

An estimate has been made by Lord Rayleigh of the least amount of energy required to produce an audible tone ; according to this it is only necessary for the air particles to vibrate through  $\frac{1}{8000000}$  cm. in order to produce an effect on the ear. The distinction easily made between notes of different quality implies in the hearer some capacity for analysing a complex sound-wave into the constituents which we have seen to exist in it. The same point is evidenced indeed by the familiar observation that when several notes are produced together, either on the same or different instruments, they may be heard separately, if listened for with sufficient attention. The analysis that the hearer performs is no doubt partly mechanical—an objective separation of the harmonic constituents made by the ear, partly psychical—a separation made by the brain in the act of perception ; and theories of audition have been mainly concerned with the share which these two processes take in the observed effect.

The mechanical part of the process of audition, the details of which were mainly worked out by Helmholtz, may be described as follows. The sound-wave, falling on the *membrana tympani*,

causes this to follow its motion ; for the membrane is very small compared with a sound-wave, and is made of material possessing considerable rigidity, so that it will not be set into independent vibrations, like a resonator, or even a simple pendulum to which a blow is given, but will move as a whole, just as the lever of a sphygmograph does under the varying pressure of an artery, or the drum of a Hürthle manometer. The *malleus incus* and *stapes* act as a chain of levers, also moving bodily, and so communicating the sound-waves still unchanged in form to the lymph behind the *fenestra ovalis* (the motion of these may be compared with that of the lever of the manometer). These levers are so arranged as to reduce the amplitude of vibration in the ratio of 3 : 2, and consequently to increase the force acting in the same ratio ; this no doubt helps to communicate the vibration in appreciable amount from air to the dense medium of the inner ear, the lymph. The perilymph, to which the motion is communicated, is bounded by practically rigid walls on all sides except at the *fenestra ovalis* and the *fenestra rotunda* ; it also therefore moves to and fro as a whole, reproducing the characteristics of the sound-wave in the air outside. Beyond this, however, it is probable that some process of selective vibration or resonance takes place. The most probable organ for this purpose is the basilar membrane ; this consists of parallel fibres, only loosely connected together, so that the membrane is tense in the direction of its breadth, but not in that of its length. In such a structure each fibre would vibrate more or less independently of its neighbours, and as the fibres vary greatly in length, and also possibly in tension, they would possess a great variety of natural periods of vibration, and would constitute a set of some thousands of resonators. These resonators would each pick out its own tone from a complex sound-wave, and communicate an impression to the corresponding nerve-ending : so that each component S. H. M. in the sound-wave would find a separate means of conveying a sensation to the brain. It is natural to conclude that this process of mechanical analysis takes place to some extent at least, although no doubt a direct action of the brain in distinguishing sounds of different characters is not to be excluded.



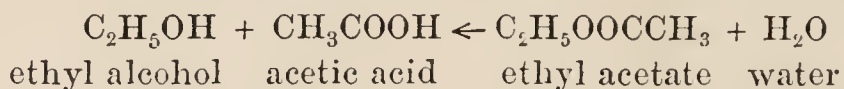
## CHAPTER VI.

### CHEMISTRY.

#### § 1. Law of mass action.

THE branch of Physics known as Chemistry has long been recognized, owing to its great extent and mass of detail, as needing independent treatment; and, apart from its detailed application to individual substances, the general principles of chemical reaction, which are of more interest from the physical point of view, have also received attention from the writers of Chemistry; but whilst such of those principles as have long been known—the theory of atomic weights, of combining proportions, of valency, and so on—are adequately treated in all books on the subject, there are others, of more recent discovery, which are not yet thoroughly incorporated in the teaching of systematic chemistry. For that reason a chapter will be devoted to the most important of them here: the subject of osmotic pressure and its bearing on solutions having been touched upon already, there remains of leading importance the law of mass action, and the phase rule—regulating between them the conditions of chemical equilibrium—and the relations of energy changes to chemical action.

A system may be called homogeneous when, whatever its chemical character, it is incapable of separation by mechanical means into dissimilar parts. Thus a mixture of gases is always homogeneous, a liquid usually so, but many instances occur of liquids which separate into two layers; e. g. ether and water: in this case each layer is homogeneous. Confining our attention for the present to a homogeneous fluid, if there be substances present capable of reacting chemically with one another, a certain state of equilibrium will be reached in which the various substances are present in calculable proportions. The conditions regulating the reaction will be more easily followed by the aid of a concrete example—one that has been much studied—the equilibrium between alcohol and acetic acid. These substances are capable of reacting as follows:—



We will call the pair of substances on the left of the equation, for brevity,



group A, that on the right, group B. Groups A and B, then, contain the same atoms, only differently arranged; and the conversion of the one group into the other belongs to the class known as *reversible reaction*, i. e. if group A be produced by mixing appropriate quantities of its components, it will gradually suffer a change, but will not be wholly converted into group B; and similarly, if we have group B to start with, it will suffer a change, but will not be completely converted into A. Thus the reaction is capable of proceeding in either sense, but in each case not completely. These facts are intended to be indicated by the symbol  $\rightleftharpoons$  which is used instead of the ordinary symbol of equality. Observation shows that in this particular case, when no further change takes place, i. e. when the system has reached its state of equilibrium, one-third of the mixture is in the form A, two-thirds in the form B. Now, this state of equilibrium may best be understood by aid of molecular considerations. Suppose a mixture A of alcohol and acetic acid made up; then the molecules of the components come into frequent collision with one another. It is only at the time of collision that such a rearrangement of atoms as is needed to produce molecules of ethyl acetate and water instead can take place; but this rearrangement does not by any means take place every time a molecule of alcohol collides with one of acetic acid: if that were the case the whole action would be over in a fraction of a second, whereas it takes hours or days to complete. We must suppose, then, that the molecules of alcohol and acetic acid possess a considerable degree of stability, and that it is only under exceptional circumstances—presumably when they collide with exceptionally great velocity—that the collision breaks up these molecules so far that they are capable of rearrangement in the form of the group B. We may put it, then, that a certain number of collisions occur every second between a molecule of alcohol and one of acetic acid, and that in a minute, but definite, fraction of these, the atoms come out of the collision rearranged, constituting a molecule of ethyl acetate and one of water. In precisely the same way, however, collisions are continually occurring between a molecule of ethyl acetate and one of water, and in a certain fraction of cases they lead to a break up of these molecules and a rearrangement in the form A. We must suppose these two processes to go on at the same time, equilibrium being reached when the amount of each is equal. This conception of a kinetic equilibrium is familiar in other departments of Physics; thus the equilibrium between a liquid and its vapour is, from the molecular point of view, regarded as such: it is not assumed that no interchange takes place between the two phases, but that some liquid molecules escape and become vapour, only that on the average an equal number of vaporous molecules become entangled in the liquid and stay there.

Now, the number of collisions of any particular kind of molecules per

second is proportional to the number of such molecules per unit of volume; so that the collisions between a molecule of alcohol and one of acid are proportional (i) to the number of alcohol molecules per c.c.; (ii) to the number of acid molecules per c.c. The number of molecules per c.c. is in any case proportional to the concentration, i. e. the number of gram-molecular-weights (or mols.) in 1 c.c. Calling this quantity  $C$  we may put—

Quantity of A converted into B per c.c. per second =  $k \times C_{\text{alcohol}} \times C_{\text{acid}}$ .  
Similarly—

Quantity of B converted into A per c.c. per second =  $k' \times C_{\text{ester}} \times C_{\text{water}}$ .  
The quantity (in mols.) of any substances suffering reaction per c.c. per second is called the *velocity of reaction*, and the constant factors  $k$   $k'$  are called the *velocity constants* for the reactions in question.

Suppose, then, only alcohol and acid to be present to start with: the reaction from left to right will go comparatively fast, that from right to left will not take place at all, for there is no ester and water to suffer conversion. As the process continues, the quantity of acid and alcohol present will constantly decrease, and so, according to the equation, the velocity of the reaction from left to right; at the same time the quantity of ester and of water present will constantly increase, and so the velocity of the reaction from right to left, until eventually the two partial reactions become equal in velocity, and equilibrium is reached. If, instead, we have only water and ethyl acetate present at first, similar changes will ensue, until the same state of equilibrium as before is reached. This state of equilibrium is evidently governed by the condition—

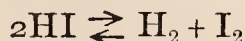
$$k C_{\text{alcohol}} \cdot C_{\text{acid}} = k' C_{\text{ester}} \cdot C_{\text{water}}$$

and the ratio  $\frac{k}{k'}$  is called the *reaction constant*: we may then write—

$$\frac{k}{k'} = K = \frac{C_{\text{ester}} \cdot C_{\text{water}}}{C_{\text{alcohol}} \cdot C_{\text{acid}}}$$

In this case the reaction constant = 4 approximately, since for equilibrium  $\frac{2}{3}$  of the mixture consists of ethyl acetate and water, i. e.  $K = \frac{\frac{2}{3} \times \frac{2}{3}}{\frac{1}{3} \times \frac{1}{3}} = 4$ ; and the value in question is practically unaltered by temperature changes. In any reaction there is a corresponding constant regulating the state of equilibrium arrived at; but in calculating the constant it must be noted that when two molecules of the same kind are involved in the reaction (alone or with others) the square of their concentration occurs in the equation for  $K$ , when three of the same kind, the cube, and so on. This is because the number of collisions involving two molecules of the same kind will be quadrupled by doubling the number of such molecules, and so on. As an example of a reaction that

has been much studied, we may take the partial decomposition of gaseous hydriodic acid that occurs at high temperatures. The reaction is—



Hence, the law of mass action gives for the velocity from left to right (decomposition of HI)  $= k C_{\text{HI}}^2$ , for that from right to left (formation of HI)  $= k' C_{\text{H}_2} C_{\text{I}_2}$ , and equilibrium ensues when—

$$\frac{k}{k'} = K = \frac{C_{\text{H}_2} \cdot C_{\text{I}_2}}{C_{\text{HI}}^2}$$

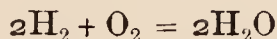
$K$  has been found to be 0.01984 at 448° cent. Substituting this number in the equation it will be found that about 22 per cent. of the acid is decomposed at that temperature. The constant is not independent of temperature in this case, as a somewhat greater fraction is decomposed at higher temperatures.

So far we have only considered mixtures of the reagents in molecular proportions, i.e. in the proportions required for the reaction. But the law of mass action gives important information as to the effect of an excess of one of the reagents. It follows from the reasoning on which the constancy of  $K$  is based, that whatever be the proportions in which the reagents are mixed, they must eventually reach concentrations such that the fraction given above is equal to the value of  $K$  for the reaction in question. Thus, reverting to the example of ethyl acetate, suppose that we have only ethyl acetate and water to start with, but that instead of these being in molecular proportion, the water is present in great excess—forming, in fact, a weak aqueous solution of ethyl acetate. Then in the equation—

$$\frac{C_{\text{ester}} \cdot C_{\text{water}}}{C_{\text{alcohol}} \cdot C_{\text{acid}}} = 4$$

the factor  $C_{\text{water}}$  will be very much larger than before; consequently, for equilibrium,  $C_{\text{ester}}$  must be much smaller. Accordingly the addition of a large excess of water causes the partial reaction in which it takes part (ester + water  $\rightarrow$  alcohol + acid) to go much further than it otherwise would, and the ethyl acetate is nearly all broken up into its components.

The reversible reaction, such as has been discussed, is really the general type of reaction, although in the usual treatment of chemistry it appears secondary. A reaction which appears to go one way only, such as the formation of water—

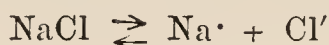


may be regarded as one in which the reaction constant is exceedingly great, i.e. the velocity of the change from left to right enormously exceeds that from right to left. If a mixture of hydrogen and oxygen be kept at 500° it will not explode, but will react slowly with formation of water—there is therefore a finite velocity of reaction; and

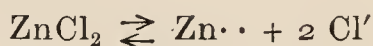


we must suppose that when two molecules of steam collide there is always a chance, though an extremely small one, that they may effect the converse change, i. e. rearrange themselves into a molecule of oxygen and two of hydrogen. This actually takes place at higher temperatures: it is observed that at  $2,000^{\circ}$  or thereabouts there is an appreciable dissociation of steam, i. e. when oxygen and hydrogen are mixed at that temperature they suffer a reversible reaction, and a state of equilibrium is reached in which some of the constituents are in the form of steam, some in the form of the elementary gases, in other words, a state similar to that of hydrogen and iodine at  $400^{\circ}$ .

One special case of chemical equilibrium is of so much importance as to deserve separate mention here. When an electrolyte, i. e. an acid, base, or salt, is dissolved in water, it suffers, according to modern views, a dissociation of a peculiar kind; it forms so-called *ions* or atomic groups possessed of positive (cation) or negative (anion) electric charges (p. 188). Thus, common salt NaCl is dissociated into one cation, viz. the sodium atom with a positive charge (indicated by a  $\cdot$ ), and one anion, viz. the chlorine atom with a negative charge (indicated by a  $'$ ); we may regard this process as a chemical reaction, and write it—



The electric charges carried by the atoms are of constant magnitude, whether positive or negative; but one ion may carry two or more charges, e. g.  $\text{ZnCl}_2$  dissociates as follows—



the zinc ion carrying twice the charge that the sodium ion did in the preceding case. Indeed the number of charges of electricity is a measure of the valency of the atoms in an electrolyte. The number of positive charges is always equal to the number of negative ones, and it is not possible to obtain a cation without an equivalent quantity of anion, or vice versa.

The chemical equations above have been written as if referring to a reversible reaction, for when an electrolyte is dissolved in water a part of it dissociates in the manner described, but a part remains in its molecular state, the electric charges neutralizing one another; and a definite state of equilibrium between the dissociated and undissociated parts is arrived at.

An acid is an electrolyte of which the cation is hydrogen. Thus, acetic acid in solution dissociates—



but the extent to which the dissociation proceeds varies very greatly from one acid to another, and consequently the amount of hydrogen ions formed; e. g. nitric acid of ordinary concentration—say 1 mol. in 10 litres of water (decinormal)—is almost completely dissociated, acetic of the same

concentration only to the extent of one or two per cent. Now, the acid properties of the liquid are due exclusively to the presence of hydrogen ions, so that the strength or *acidity* of an acid depends solely on the extent to which it is dissociated into ions.

Similarly, a base is an electrolyte of which the anion is hydroxyl; e.g. baryta in solution gives—



and the basic properties depend solely upon the amount of hydroxyl ions present.

In all cases the dissociation is favoured by dilution, so that when a salt, acid, or base is dissolved in a very great amount of water it may be regarded as completely broken up into its ions.

Water itself dissociates to a small extent, yielding both hydrogen and hydroxyl ions—



but the amount of them is very small, only about 1 part in 10,000,000,000 being dissociated at ordinary temperature. It may, however, on this account, be looked upon either as a very weak acid or a very weak base.

The velocity of all reactions is increased by rise of temperature, and that very rapidly; the rates of increase per  $1^\circ$  in cases that have been measured vary from 2 % to 13 %, and are mostly nearer the higher than the lower limit. This increase is in geometrical, not arithmetical progression, i. e. each degree raises the velocity by the same fraction of its then value, not by the same actual amount as in the expansion of gas. Hence, the increase over a wide range of temperature is enormous: if the rate per  $1^\circ$  is 10 %, that for  $100^\circ$  is to be found by reckoning compound interest on the amount at  $0^\circ$  one hundred times over at 10 %. Thus, the rate of decomposition of dibromsuccinic acid in aqueous solution has actually been measured at  $15^\circ$  and at  $101^\circ$ , and proves to be 3,000 times as rapid at the latter temperature as at the former. In consequence of this the range of temperature over which any particular reaction is observable is very restricted, e.g. the combination of oxygen and hydrogen goes on at a moderate and measurable rate at the boiling-point of sulphur ( $448^\circ$ ); but at atmospheric temperature, if it takes place at all, the reaction is so slow that it is not possible to observe it; whilst at  $700^\circ$  or thereabouts it is so rapid as to constitute an explosion.

In a reversible reaction both partial reactions are accelerated by rise of temperature, but not necessarily at the same rate. If they are the state of equilibrium will not be affected; this is the case for the formation of ethyl acetate. But usually one reaction gains on the other as temperature rises; thus, the decomposition  $2\text{HI} \rightarrow \text{H}_2 + \text{I}_2$  gains on the formation  $\text{H}_2 + \text{I}_2 \rightarrow 2\text{HI}$ , and consequently at high temperatures more hydriodic acid is decomposed than at low.



Certain changes in organisms have been studied for the influence of temperature, and appear to behave like chemical reactions; thus, the respiration of plants has been found to increase nearly 10% per 1° between 0° and 25°, and seems to go on increasing (not reaching a true optimum point) until the excessive rate of action destroys the plant. The action of enzymes has been found to increase at about the same rate, and the excessive action of the human organism during fever may probably be classed with these effects.

The changes occurring in the fluids of the digestive organs, the blood, &c., are essentially chemical changes, and consequently follow essentially the same laws as ordinary chemical reactions that have been followed out in the laboratory.

## § 2. The phase rule.

In contradistinction to the definition given at the beginning of the last section, a system is called heterogeneous when it can be separated mechanically into parts; and each of the parts is called a *phase*. Thus, if water and its vapour be present together, as in the boiler of a steam engine, we have two phases; if a solid salt lie at the bottom of a vessel containing its aqueous solution, we have again two; and if, further, some water vapour exist above the liquid there are three phases to consider. Generally gases form only one phase, since they all mix together; liquids may form one or two or even more according to their miscibility, e.g. if phenol and water be mixed in equal proportions they will form two layers, one of phenol, which however contains some dissolved water, and one of water, which likewise contains some dissolved phenol: by warming the mixture to 60° or 70° the two liquids come to dissolve more and more of each other till they completely mix, and then there is only one phase. Solids, on the other hand, usually constitute a phase each, since they do not mix at all.

The number of possible phases in a system is associated with the number of chemical *components* that go to form it. In reckoning these it is not necessary to push the analysis as far as it will go; it is sufficient to find what is the least number of components by which the analytical composition may be expressed: the meaning of this will become clear on considering some particular cases. Thus, if we have ice, water, and steam together (which is possible under certain conditions), there is no need to take into account the fact that these substances are formed by the combination of oxygen and hydrogen, nor to enter into discussions as to the molecular weight of the water in its various phases, or its electrolytic dissociation. It is sufficient to note that the analytical composition of all three is identical, and to say that there is only one component. Suppose, however, some hydrochloric acid added to the water; then the



phases will not all have the same composition, for the ice will be unaffected, remaining pure  $\text{H}_2\text{O}$ , the liquid phase will contain a certain percentage of  $\text{HCl}$ , and the vapour will also contain a certain percentage of  $\text{HCl}$ , but not in general the same as the liquid. However, to express the composition of these three phases, it is not necessary to write down the percentage of hydrogen, oxygen, and chlorine in each (three substances), for the oxygen is throughout combined with a fixed proportion of hydrogen, and the chlorine with another fixed proportion of hydrogen; hence, we shall know all about the composition by stating the percentage of water and hydrochloric acid in each phase (two substances). The number of components therefore is two.

Now, considering first, for simplicity, a system of one component, say  $\text{H}_2\text{O}$ , let us find in how many ways it may vary independently. For this purpose we will start by supposing that there is only one phase, steam. The steam can obviously not be varied in composition, but it may be varied in pressure while the temperature is kept constant (provided we do not increase its pressure so far as to make it condense), and it may be varied in temperature while kept under constant pressure (with the same proviso as to condensation brought about by cooling). There are, consequently, two independent variations possible without passing out of the given phase, or, as it is stated, two *degrees of freedom*. We have here left out of account a mere change in the quantity of steam, apart from any change in its condition, for, as we shall see throughout, the quantity of any phase present is without influence on the state of equilibrium reached (in marked contrast to the effect of the quantity of the different constituents within one phase, as considered in the preceding section). But suppose that the steam is brought into a state in which some of it condenses to water: there are then two phases. We have seen already (p. 86) that steam in order to be saturated, i. e. to exist in equilibrium along with liquid water, at any given temperature, must be at a definite pressure corresponding to that temperature—known as the saturation pressure. It is, consequently, not now possible to effect two independent variations in the state of the system without the disappearance of one of the phases; for if the temperature of the system be raised the pressure must be raised too, or it will all be converted into vapour, and if the temperature be lowered, the pressure must be lowered too, or it will all be condensed into liquid. The change in pressure is therefore no longer independent of the change in temperature, so that there is only one degree of freedom. Next suppose the mixture of water and steam to be cooled, taking care to adjust the pressures so that neither phase disappears, till ice begins to form; there are now three phases. It will be found that there are no degrees of freedom left, i. e. no change is possible in either temperature or pressure without one of the phases disappearing; if the temperature is raised the ice will melt, if it is

lowered the water will all freeze, unless that is prevented by an appropriate rise of pressure, and then the vapour phase would disappear. The condition of temperature and pressure under which all three phases can coexist is called the triple point. It is slightly above  $0^{\circ}$  in temperature and about 5 mm. in pressure.

These relations are perfectly general; the appearance of a new phase always implies the loss of a degree of freedom, so that if the system be of one component only, it has two degrees of freedom when one phase is present, one when two phases are present, forms a triple point when three are present, and can in no case contain more than three phases. There are instances in which four or more phases of a single substance are known, e. g. sulphur in the forms of (i) vapour, (ii) liquid, (iii) monosymmetric crystals, (iv) rhombic crystals; but it is impossible to have more than three of these in equilibrium together.

The triple point and associated phenomena may be clearly represented by a diagram with temperature and pressure as ordinates, such as has already been used to explain the properties of the melting-point (Fig. 46). The lines on that figure have already been referred to, with the exception of *DE*; this is the boundary between the solid and liquid states, in other words the melting-point line. *D*, as already explained, is the melting-point of the solid when under the pressure of its own vapour; it is consequently the triple point, for the solid and liquid can there not only exist alongside of one another, but in presence of their vapour too: for benzene, to which Fig. 46 primarily refers, the triple point is at about 36 mm. pressure and  $5^{\circ}.58$ . Suppose then a mixture of solid, liquid, and vaporous benzene kept in a cylinder closed by a piston; the pressure being adjusted to the value just mentioned and the correct temperature being maintained by a bath; now push in the piston. The pressure will not at first rise, for as the volume gets less, more and more vapour will condense; when no vapour is left, however, and there are consequently only two phases present, the system will possess one degree of freedom—it can be changed in pressure, but only on condition that the temperature changes in a definite manner too, if both phases are to be preserved. This definite condition of change is shown by the line *DE*; in the case of benzene, as the pressure rises, the temperature must rise too, i. e. the melting-point is raised by pressure. This is because solid benzene is a little denser than liquid, and consequently the increasing pressure would squeeze the whole substance into the solid form, were it not counteracted by a rise in temperature. The line *DE* accordingly slopes to the right, upwards; the slope is very slight however (much exaggerated in the figure), amounting only to one- or two-hundredths of a degree per atmosphere of increased pressure: consequently the melting-point under atmospheric pressure is practically identical in temperature with the triple point. For water the line *DE* slopes a little to the left,



because ice is less dense than water; pressure consequently tends to convert it all into the liquid form, and has to be counteracted by lowering of temperature.

From the triple point *D* therefore there start out three lines, *DB*, *DC*, *DE*, dividing the diagram into three areas, *CDE*, *EDB*, *EDC*. The point *D* represents the simultaneous existence of three phases; each line two phases; each area one phase only. Further it is easily seen that the point allows of no movement at all, each line allows of movement in one direction (*viz.* along its length), in other words one degree of freedom; while within the spaces, independent movement in two directions is possible, *i.e.* changes in pressure and temperature can be made separately.

If a system has two components, *e.g.* water and hydrochloric acid, then it may possess three degrees of freedom, for in addition to temperature and pressure, we may vary the percentage, say, of hydrochloric acid in it. Again, by introducing one new phase, one degree of freedom is lost, so that starting from a single phase we may add three more, and reach a maximum of four. Similarly if there be three components, five phases may coexist, and so on. The *phase rule* which is due to Gibbs may then be stated generally as follows:—‘The number of degrees of freedom of a chemical system is equal to the number of components increased by two, and diminished by the number of phases.’

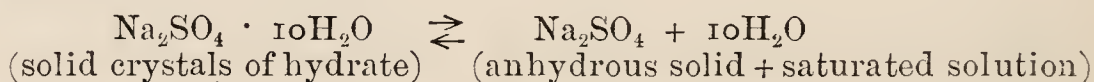
The melting-point is really only one of a class called *transition points* at which one phase passes over into another. One of the most familiar instances of such is the transition between the two crystalline forms of sulphur; the rhombic form, in which that substance is usually found, is stable at ordinary temperatures, but at  $95^{\circ}$  it is converted into another form, that of crystals belonging to the monosymmetric system. The conversion is a much slower process than that of fusion, but if rhombic crystals be kept at  $100^{\circ}$  for some hours they will be found partially converted into the other form, and the conversion will proceed steadily until the one form is completely replaced by the other. If then the product is maintained a long time at  $90^{\circ}$  it will suffer the converse change, and with equal completeness. The transition point is therefore a point of interchange of stability, like the melting-point: and when the form stable at lower temperatures (rhombic) passes over into that stable at higher temperatures (monosymmetric) there is always an absorption of heat, the *latent heat of transition*, exactly comparable with the latent heat of fusion. In the case of sulphur the transition is accompanied by an expansion, but that is not necessarily the case; the rule as to absorption of heat is universal.

Another example is the transition between the ordinary form of iron, which is magnetic, at about  $700^{\circ}$  into another form which possesses no magnetic properties. Again, it has lately been found by Cohen that



ordinary tin is not stable below  $20^{\circ}$ , but is converted into a grey crystalline form; this sometimes takes place spontaneously in very cold winters, a sort of 'tin-pest' breaking out on the surface of the metal, and gradually spreading. Ordinary tin can be 'infected' by means of specimens of the grey form, and thereby converted, more or less rapidly, into the latter form, which is the more stable at low temperatures.

Innumerable cases of transition amongst compounds are known; e.g. Glauber's salt melts in its own water of crystallization at  $32^{\circ}$ . This process is really the transition:



All transition points, like the melting-point, are but little affected by pressure, because the change in volume accompanying the transition is small; accordingly it makes little difference whether the temperature of transition be measured at atmospheric pressure, or at the saturation pressure of the substance concerned. What little difference there is, however, follows the same rules as in the case of the melting-point.

The two leading results as to transition may then be stated formally as follows:—

'The system which is formed from the other with absorption of heat, is the more stable at temperatures above the transition point, and vice versa.'

'The transition temperature is raised or lowered by pressure according as the system expands or contracts in passing from the form stable below the transition point to that stable above it.'

The importance of the phase rule as a guiding principle among the phenomena of chemical equilibrium may be illustrated by the following (out of very many possible) examples. Amongst drying agents, anhydrous copper sulphate and strong sulphuric acid are commonly used: their action however is different in principle. Copper sulphate is usually obtained in crystals containing five molecules of water; these on drying (either by heating or in a vacuum desiccator) dissociate, leaving a salt with three molecules of water; on further drying this breaks up, yielding a monohydrate, and the latter finally dissociates in a similar manner, giving the anhydrous salt. We need only consider the last reaction, which may be written—



Now if some of this hydrate be placed in a tube and the pump applied to evacuate it, so that it gives off some water, we have a system consisting of three phases, viz. two solids, the hydrate, and the anhydrous salt, and one gas, steam. The system is made up of two components, clearly, for it is only necessary to analyse any one of the phases into its percentage of  $\text{CuSO}_4$  and of  $\text{H}_2\text{O}$  to state its composition. According to

the phase rule, therefore, the degrees of freedom possessed by the system amount to  $2 + 2 - 3 = 1$ . This means that if the temperature be changed the pressure must be changed in a definite manner: there is a fixed pressure (called the dissociation or saturation pressure) corresponding to each temperature. The case is in fact similar to that of a liquid and its vapour. If a vessel containing water and water vapour, kept at say  $15^{\circ}$  cent., be connected to a pump, when the pressure is reduced to about 13 mm. the water will boil, and it will not be possible to reduce the pressure further, so long as any water is left, for if water vapour be pumped out, it will merely be replaced by further evaporation; and if steam at higher pressure be let in, it will condense. In the same way the hydrated salt  $\text{CuSO}_4 \cdot \text{H}_2\text{O}$  possesses at  $15^{\circ}$  a definite dissociation pressure of about 0.6 mm. Consequently if the pressure be reduced below this point it will dissociate and give off water vapour, while on the other hand if damp air be passed over the anhydrous salt, the reaction will proceed in the opposite sense, and water will be removed from the air till the pressure (of water vapour) has fallen to 0.6 mm.; hence anhydrous copper sulphate is a fairly effective drying agent, removing some  $\frac{1}{2}\%$  of the moisture in the air, but it is incapable of going beyond that point.

The drying action of strong sulphuric acid is different: sulphuric acid being a liquid, it, with the water it absorbs, constitutes only one phase, and the vapour over it another. Now in a system of two components with two phases there are  $2 + 2 - 2 = 2$  degrees of freedom; i.e. there is not in this case a definite pressure corresponding to any particular temperature; but, the temperature being kept constant, it is still possible to vary the vapour pressure by varying the composition of the acid; e.g. if a bulb containing 10 gms. of pure  $\text{H}_2\text{SO}_4$  be used and some air containing 0.1 gm. of moisture be passed over it, most of this will be absorbed, and the strength of the sulphuric acid thereby reduced to about 99%; such acid has a vapour pressure say  $x$ , and consequently will not absorb more moisture from the air passed through it if the pressure of water vapour in the air be lower than  $x$ . If however another bulb be put behind it, containing another charge of pure  $\text{H}_2\text{SO}_4$ , this will absorb more moisture from the air which has passed through the first: thus the drying power depends essentially on the quantity of sulphuric acid used, whereas with copper sulphate the quantity is immaterial, so long as there is any of the anhydrous substance left.

Another example of the phase rule may be found in the aëration of the blood.

Here the equilibrium is between the two forms of haemoglobin, oxidized and reduced, and the oxygen of the air; there are thus two solids and a gas, and the case is similar to that of copper sulphate and water vapour. The oxyhaemoglobin gives off oxygen—dissociates—at a definite pressure, so that it can be alternately formed and dissociated by



exposure to oxygen at a comparatively high pressure (in free air) and at low pressure (in the tissues). This at least is the general outline of the phenomenon: it is really more complicated and less perfectly understood than the dissociation of a hydrated salt.

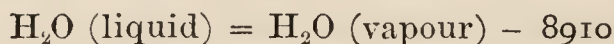
### § 3. Thermochemistry.

Changes of state, both physical and chemical, are in general accompanied either by an absorption or evolution of heat; we have seen, for instance, that when ice is converted into water or water into steam, a considerable absorption of heat takes place; again, that any transition, such as that from rhombic into monosymmetric sulphur, involves absorption of heat; while it is a well-known fact that many chemical processes involve a large solution of heat, e.g. when carbon combines with oxygen of the air and burns, when acid is mixed with alkali, and so on. Examining more closely these thermal changes, we find that in accordance with the law of conservation of energy, a distinction must be made between cases in which the substance suffering the change remains of the same volume as before or not. As there is no essential difference between physical and chemical instances we will take one of the former, for simplicity.

When a gram of water at  $100^{\circ}$  is boiled, some 536 calories of heat have to be put into it. This implies that the steam formed contains a very much larger amount of energy than the water it was formed from; at the same time, not all the heat absorbed is spent in adding to the stock of energy in the substance; for whilst the water occupies only 1 c.c., the steam occupies about 1,654 c.c.; therefore during the process of evaporation the atmospheric pressure has been overcome through a volume equal to the difference of these, and work has been done by the expanding steam to the extent  $= (1654 - 1) 1013200$  ergs, i. e. the increase in volume multiplied by the pressure. In calories this is  $\frac{1653 \times 1013200}{42000000} = 41$ .

Since then the substance we are dealing with gives out during the process mechanical work equivalent to 41 calories, while it absorbs 536, there can only remain  $536 - 41 = 495$  as the *increase in internal energy* of the water on conversion into steam. This is sometimes called the internal latent heat.

For chemical purposes it is more convenient to refer results to a gram-molecule of material than to a gram; the molecular weight of steam being 18, we may put the molecular internal latent heat  $= 495 \times 18 = 8910$  cal., and to express the heat change it is customary to write down the chemical equation with the amount of heat *evolved* in it at the end. The reaction we have just considered is then—



the negative sign before the quantity of heat implying that heat is



*absorbed* during the change indicated. In each case it should be carefully specified whether the internal or total heat change is meant. If however the conversion is from one solid or liquid into another, the change in volume is so small that the external work done may be neglected by comparison with the absorption of heat, and the distinction ceases to be of consequence. Thus ice in melting absorbs 80 calories per gram, while the amount of work done (in this case done on it by the atmosphere, since the ice contracts) does not amount to the equivalent of  $\frac{1}{100}$  calorie.

Turning now to a chemical case we note precisely the same distinctions. If only solids and liquids are concerned we may ignore any changes in volume; thus iron and sulphur combine with evolution of heat, the reaction only needing to be started by a local high temperature; it may be written—



This statement implies evolution of heat (since the sign is +), and that when 56 grams of iron combine with 32 of sulphur 23,800 calories of heat are given off. In other words, the *internal energy* possessed by 88 grams of ferrous sulphide is *less* by 23,800 calories than that of the elements it was formed from. This quantity is called the *heat of reaction*.

If, on the contrary, gases are concerned, the external work done by the reacting substance if it expand, or done on it if it contract, is considerable enough to take into account. Practically a reaction is performed either under the constant pressure of the atmosphere, free contraction or expansion being allowed, or else in a space of constant volume (such as Berthelot's calorimetric bomb). In the latter case evidently no external work is done, and the true or internal heat of reaction is directly measured. When the volume is allowed to change the circumstances are slightly more complicated. As example we will suppose four grams of hydrogen to be burnt under ordinary conditions. It is then observed that 136,720 calories are evolved: the equation may be written—

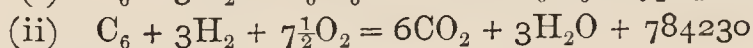


the water formed being collected as liquid. Now three mols. of gas disappear in the equation, while only a liquid is formed, so that work is done by the atmospheric pressure on the system—e.g. if the combustion takes place at the top of a eudiometer tube, mercury is driven in by the atmospheric pressure to take the place of the gas formerly there. The work done in this way can easily be calculated, as in the preceding example, and since a mol. of any gas occupies the same space when under the same conditions of temperature and pressure, the same amount of work is done for each mol., the amount being (at atmospheric temperature) about the equivalent of 580 cals. Hence  $3 \times 580 = 1740$  cals. of heat is produced by means of this work, and of the observed generation of heat so much is accounted for, and there remains only  $136720 - 1740 = 134980$  cals. really due to the combination. This is the

heat that would be evolved if the combustion took place in a closed space such as the calorimetric bomb, when no question of external work comes in; it is therefore the true decrease in energy of the materials: four grams of hydrogen and 32 of oxygen uncombined possess 134,980 cal. more energy than 36 grams of water.

It follows from the law of the conservation of energy, that if a chemical system be transformed from a certain initial state *A* to a certain final state *B*, it gains or loses a fixed amount of energy, however the transformation be effected; for if we transformed it by a certain process and it gained energy, and we then transform it back by a different process, it must lose precisely the same amount of energy, otherwise it would not in the end have the same content of energy as at starting, but being in the same state at the end as at the beginning it must have the same content of energy.

The important results that this principle leads to may best be understood by the aid of an example. Suppose it is desired to know how much heat is evolved or absorbed in forming benzene ( $C_6H_6$ ) from its elements. It is not possible to synthesize benzene directly, so to determine the amount of heat we must have recourse to indirect methods. Now carbon and hydrogen can be burnt in a calorimeter, and the heat evolved measured; so too can benzene; and it is found that the heat of combustion of six grams of hydrogen is 202,470 cal., of 72 grams of carbon 581,760, making a total of 784,230, while 78 grams of benzene on combustion give out 797,900 cal. Hence if we could synthesize benzene from carbon and hydrogen, and the heat evolved in doing so be  $x$ , we may compare the following processes:—



(i) and (ii) lead from the same initial to the same final conditions, so that the total change in energy must be the same in each; it follows that  $x + 797900 = 784230$ , or  $x = -13670$ . Accordingly the requisite quantities of carbon and hydrogen on combining to form benzene *gain* 13,670 cal. of energy. Fig. 55 may serve as a simple illustration of the principle here involved. *A* represents the initial state of elementary carbon, hydrogen, and oxygen; *B* the final state of carbon dioxide and water; in passing from *A* to *B* the energy-content falls by 784,230 cal. *c* represents the intermediate condition of benzene and oxygen, and it is found that in passing from *c* to *B* the energy falls by 797,900 cal.; hence it is obvious that *c* is 13,670 cal. higher in energy than *A*.

*B* — In all the above reasoning it must be carefully remembered that we are concerned with changes of internal energy only, not with work done by the reacting system on other bodies, so that the true heat of reaction, i.e. at constant volume, must always be employed.



A reaction such as the formation of benzene from its elements, in which the heat of reaction is negative, is called *endothermic*: the more frequent class in which the heat of reaction is positive is called *exothermic*.

A convenient abbreviated notation for heats of reaction is to write within square brackets the substances out of which the compound is formed with a comma between each, and put the heat evolved on the opposite side, e. g.—

$$[\text{C}_6, \text{H}_6] = - 13670$$

$$[\text{H}, \text{Cl}, \text{Aq}] = + 39300$$

The symbol Aq represents an indefinite quantity of water, used as a solvent,  $\text{H}_2\text{O}$  being retained for the molecular quantity, such as enters into reaction; consequently the latter equation means that when one gram of hydrogen and 35.4 grams of chlorine, both in the gaseous state, combine to form hydrochloric acid, and dissolve in water, 39,300 calories are evolved. This is a strongly exothermic reaction; as a rule the reactions which take place most readily are those in which a good deal of heat is liberated; but for more exact information as to the bearing of heat of reaction on chemical equilibrium, see below.

The most important thermochemical data, so far as medicine is concerned, are the heats of oxidation of the elements forming organic compounds. These are—

$[\text{H}_2, \text{O}] = + 67490$	or hydrogen	.	.	33745	cal. per gm.
$[\text{C}, \text{O}] = + 26300$	carbon to monoxide	.	2192	„	„
$[\text{C}, \text{O}_2] = + 94640$	carbon to dioxide	.	7887	„	„
$[\text{S}, \text{O}_2] = + 71000$	sulphur	.	.	2219	„

Nitrogen, in the decomposition of organic compounds, is usually given off in the elementary state.

The thermal efficacy of food-stuffs (and coal) may be roughly calculated by means of the above numbers; only roughly, for we have seen, in the instance of benzene, that the energy of the elements when combined is not quite the same as when free. In reckoning the thermal value in this way, if there is oxygen in the compound it should be regarded as combined with the hydrogen, and so reducing from the total supply available. As an instance of the mode of calculation we will take a proteid of composition C 53 %, H 7 %, N 16 %, O 22 %, S 1 %, ash 1 %. Then in 1 gram we have 0.22 gm. of oxygen: this may be regarded as combined with  $\frac{1}{8}$  of its weight in hydrogen, i. e. 0.0275 gm.; deducting this from the 0.07 gm. of hydrogen we have 0.0425 left for combustion. The heat generated will therefore be, approximately—

0.53	gm. of carbon	@	7887	=	4179	calories
0.0425	„	hydrogen	@	33745	=	1434 „
0.01	„	sulphur	@	2219	=	22 „
						<hr/>
						5635 „



An estimate of average human diet gives the following as the total supply of energy to the body per day :—

100 gm. proteid	= 550000 cal.
100 „ fat	= 950000 „
240 „ carbohydrate	= 960000 „
	<hr/>
	2460000 „

This in mechanical units is  $2460000 \times 4.2 = 10,330,000$  joules per day = 120 joules per sec. (watts). This is the average total activity of a full-grown man throughout the day and night: the average of mechanical work done by those engaged in muscular occupations may be taken at 20 watts (or say 48 per hour during ten working hours). Hence the efficiency of the human machine, averaged throughout the twenty-four hours, may be taken as one-sixth.

#### § 4. Heat and chemical equilibrium.

In Mechanics one is familiar with the principle that the *potential energy of a system tends to run down to a minimum*; this was partly taken into account in considering the various kinds of equilibrium on p. 31. But as the principle is applicable throughout Physics, and has a special bearing on chemical equilibrium, it will be reconsidered here. Beginning with a simple mechanical example, suppose a ball to be placed on an incline; it has a certain amount of potential energy, due to its height

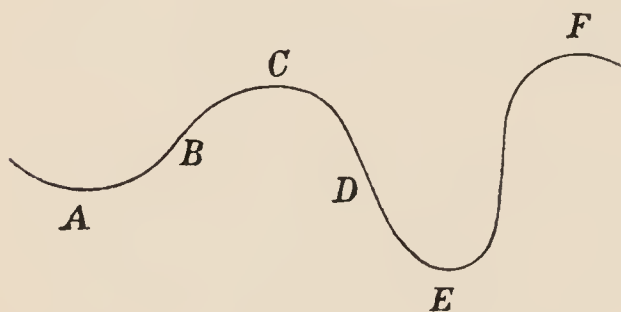


FIG. 56.

above the earth, the amount being measured by its weight multiplied by the height above whatever is chosen as the standard level. If the ball be moved higher up it will therefore possess more energy; if lower, less; and accordingly if left to itself it will roll downwards, i.e. in the direction of

diminishing potential energy. If it be at the bottom of an incline, say at *A* (Fig. 56), it will be in equilibrium, and in particular in stable equilibrium; this is because the level rises in all possible directions from *A*, and such a point is called a point of *minimum* level. This does not mean that it is the lowest point of all, for *E* is lower; but merely the fact that the level rises in all directions from *A*. Again, if the ball be at the top of an incline, as at *C*, it is in equilibrium, but this time unstable; its height, and therefore its potential energy, is a *maximum*, i.e. the level slopes down from that point in all directions, although *C* is not the highest point of all, for *F* is higher. We see, then, that

stable equilibrium corresponds to a minimum of potential energy, unstable to a maximum; while a point which is neither a minimum nor a maximum, like B and D, is not a point of equilibrium at all. Further, the equilibrium at E is more stable than that at A, nevertheless it is not possible to transform A into E without giving the body sufficient energy to climb the intervening maximum. The term *metastable* is sometimes used to distinguish a condition like A, which is stable, and yet not the most stable condition possible.

A closer analogy to the chemical case is supplied by the action of a siphon. Suppose A and E connected through C by a siphon; then water could not flow of its own accord from A to C, but if the siphon were once started it would enable an indefinite quantity of water to pass from the metastable condition A to the stable E; and in order to start the siphon it is only necessary to impart a small impulse from without, just to raise enough water to fill the siphon.

Applying these considerations to the case of chemical reaction, we may say that a mixture of benzene vapour and oxygen is a system in metastable equilibrium: it possesses a very considerable degree of stability, and indeed may be kept for an indefinite time without change; yet it is not so stable as the water and carbon dioxide that may be formed from it, only an intermediate condition of greater potential energy has to be passed through to transform the one into the other. This can be accomplished by means of local heating—say by passing an electric spark through some point of the mixture; this causes an explosion, and the whole is transformed into the stable form of water vapour and carbon dioxide; we may pass sparks repeatedly through this without getting it back into the metastable form again.

A similar process constitutes the metabolism of the animal organism. The food-stuffs possess a considerable degree of stability—sugar, starch, fat, and so on, can be kept indefinitely; but they possess a large amount of potential energy that can be given out by combining them with oxygen of the air to form carbon dioxide and water. This is continuously accomplished in the organism by a mechanism that builds up the materials into the form of protoplasm, which if not actually unstable, no doubt constitutes a metastable form possessing so much energy that it easily breaks down; the breaking down of this by oxidation then supplies the energy by which the organism is kept in action.

It was at one time thought that the potential energy of a chemical system, according to which its transformations are determined, could be measured by means of heats of reaction. It is true indeed that by means of the heat evolved in a reaction we measure the change in the *total energy* of the reacting system; but it does not follow that this total energy is what must be taken into account to decide upon the stability or instability. As a matter of fact it is a part only of the energy, which is called



the *free energy*, that determines that. To continue the mechanical analogy, we may suppose the ball in Fig. 56 to be set spinning about a vertical axis : then it possesses energy of two kinds, the potential energy we have already considered, and kinetic energy of rotation ; now the ball might be spinning at A so rapidly as to possess more energy on the whole than if it were at rest at c. Yet, though the circumstances would be more complex than in the previous state, it would still be true that A was the position of stable equilibrium, and the ball would not pass of its own accord from A to c ; it is therefore in this case only a part of the total energy which is decisive as to the equilibrium of the ball. With regard to the chemical case, something similar is true ; though it is not possible within the scope of this book to explain fully the method of determining this part—the so-called *free energy*, the conclusion is that the change of total energy, as measured by heat of reaction, is an unsafe guide. True, in the majority of cases a reaction will proceed in the sense that causes heat to be evolved ; but the opposite case, that of an endothermic reaction, is by no means rare ; thus among processes of a more physical character the formation of a freezing mixture out of salt and snow is a striking instance ; the mixture takes place of its own accord, although accompanied by so much absorption of heat that the product falls many degrees below zero. Again, to take a purely chemical example, hydrochloric acid, being a stronger acid than hydrofluoric, decomposes sodium fluoride in aqueous solution, although some 2,300 calories are absorbed in the process.

The rule then is that a reaction will proceed so that the free energy of the system is reduced to a minimum. In some cases changes of free energy can be determined by means of electromotive force ; but unfortunately this method is not general, as most substances are not electrolytes, and in the majority of reactions the changes in free energy are as yet unknown. Nevertheless certain deductions from this principle can be made.

In the first place, the value of a food-stuff in supplying energy to the organism is not to be reckoned solely by its heat of combustion ; this, as we have seen, is a rough guide, showing in fact what is the total energy the substance can give up when converted by the processes of the organism into water, carbon dioxide (and urea) ; but it is not this total energy which we are concerned with. In a steam engine it would be, and if the body were a heat engine it would be so in that ; in a heat engine all the chemical energy of the fuel is converted into heat, by burning it, and that heat used for the production of work (by means of steam or otherwise). Now a heat engine can only convert into work at most a definite fraction of the heat it consumes, viz.  $\frac{(T_1 - T_2)}{T_1}$ , where  $T_1$  is the absolute temperature at which the heat is supplied, and  $T_2$  that at



which the waste is withdrawn (p. 64). If the body were a heat engine it could only work between its own temperature—say  $37^{\circ}$  cent.—and that of the surrounding air, say on the average  $15^{\circ}$  cent. The efficiency of such an engine could at most be only—

$$\frac{37-15}{273+37} = 7\%$$

whereas the efficiency of the human body during actual work has been found to be about 26.5 % (p. 61). Hence the mechanism of production of work in the organism must be quite different: the energy of food is not first transformed into heat and then used, but is transformed directly into mechanical energy; it is therefore incorrect to speak of food as the fuel of the animal body—the metaphor is no doubt strengthened by the observation that much heat is given off from the organism in its working, but the analogy between it and a heat engine is really misleading, the two mechanisms being fundamentally different. Consequently in order to estimate the value of a food-stuff as a means of producing the work necessary to the organism, it would be necessary to know the amount of its free energy; but that quantity has not yet been determined.

In the second place, it has been shown (by van 't Hoff) as a deduction from the principle of free energy that in any reversible reaction by which one system of bodies is converted into another—

*‘Rise of temperature favours the system formed with absorption of heat.’*

We have already seen that this principle is applicable to cases of transition from one phase to another. To take the simplest physical instance, water and steam are two phases that can be reversibly converted into one another, and steam is formed out of water with absorption of heat; consequently rise of temperature favours the formation of steam, fall of temperature, water. This is in fact only another aspect of the principle stated at the end of § 2, that in passing from the form stable at low temperatures to that stable at high temperatures, heat is absorbed.

But van 't Hoff's statement was intended to apply rather to equilibrium in homogeneous systems. This, as we saw, follows the law of mass action, and when any reversible reaction  $A \rightleftharpoons B$  can take place a certain state of balance is arrived at, in which some of  $A$  and some of  $B$  are present; such an equilibrium however holds only for a particular temperature; when the temperature is altered the equilibrium will be displaced, the percentage of  $A$  being increased or decreased. The rule stated above shows which way the displacement will take place; thus carbon dioxide is formed from carbon monoxide and oxygen with a large evolution of heat. At low temperatures the reaction is carried out with practical completeness in one sense—nothing but carbon dioxide is formed when a mixture of the gases is sparked or otherwise made to react. But rise

of temperature favours the system formed with absorption of heat, i.e. the carbon monoxide and oxygen, so that at  $2,000^{\circ}$  the reaction  $2\text{CO} + \text{O}_2 \rightleftharpoons 2\text{CO}_2$  is decidedly a reversible one, quite a considerable amount of the system on the left of the equation existing in equilibrium; this has been noted in the gases of a blast furnace—the carbon dioxide is said to be partly dissociated at so high a temperature. Conversely, as we have seen that hydriodic acid is dissociated to a somewhat greater extent at high temperatures than at low, we may conclude that the system  $\text{H}_2 + \text{I}_2$  is formed from the system  $2\text{HI}$  with absorption of heat.

Now the atmospheric temperature is a low one, being only about  $290^{\circ}$  above the absolute zero, whereas we have to deal with temperatures up to  $2,000^{\circ}$ , and even higher. Hence it is natural that in most cases of equilibrium at low temperatures the system formed with evolution of heat should prevail; and it is found in practice that reactions which take place energetically at ordinary temperature are accompanied by a large generation of heat; this led at first to the opinion that chemical reaction necessarily took place in the direction to evolve heat. That opinion is not correct, since endothermic reactions also are known; but at the absolute zero it would be true.

## CHAPTER VII.

### ELECTRIC CURRENTS.

#### § I. General properties of currents.

AN electric current is a phenomenon to the production of which two things are necessary : (i) a source of electrical energy ; (ii) a conducting circuit. What the current really is, in itself, is by no means completely known at present, nevertheless the properties it possesses and actions it exerts have been investigated with a thoroughness and exactitude that are not excelled in any other department of Physics. What we have to do, therefore, is to describe these properties, and in order to do so we shall begin by leaving out of consideration the source of electrical energy, which may be a voltaic cell, a dynamo, &c., and confining attention for the moment to the effects observed in and near the conducting current.

To begin with, then, substances may be classed into those which do and those which do not conduct electric current, or to be more accurate, may be classed according to the facility they possess for conducting it. In the first class come all the metals ; these are good conductors : in the second class a group of bodies whose relation to the electric current is peculiar, since they suffer chemical decomposition when the current is led through them ; these are the so-called electrolytes, and include salts, acids, and bases, in aqueous solution, and also, in many cases, when fused. A third class may be constituted of all the remaining bodies, their conducting power being extremely small, and in the case of gases perhaps nil ; amongst these bodies, called insulators, some that are important for electrical purposes, from the very fact that they strongly resist the flow of current through them, are mica, glass, ebonite, paraffin.



Next, in order to allow an electric current to flow, it is not sufficient to provide a path of conductors, whether metallic or electrolytic, but the path must be a completely closed circuit. For it is found that electric currents flow only in closed, circuital paths; any break in the chain of conductors, made by interposing an air space or a piece of ebonite, &c., stops the current.

Hence it is easy to construct *electrical keys or switches*, by means of which the current may be turned on or off as desired. Such a key is shown in Fig. 57; it consists of two brass blocks,

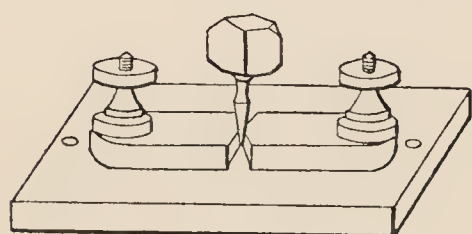


FIG. 57.

mounted on an ebonite base. Each block carries a 'binding screw' for conveniently making metallic connexion by means of wires with other pieces of apparatus. The space between the blocks can be filled up at will by a tightly fitting conical brass plug, provided with an ebonite

handle. When the plug is inserted the current flows freely through the apparatus, when it is removed the metallic circuit is broken and the current stops. Many other patterns of keys are in use.

The principal effects produced by an electric current are (i) thermal, (ii) magnetic, (iii) chemical; we will consider these briefly, in the order stated.

(i) When an electric current flows through any substance whatever it meets with some resistance, and in overcoming this it generates heat; the energy of the current is in fact converted into heat—a process of dissipation analogous to that which friction causes in the case of ordinary mechanical motion. Now it will be necessary, as we shall see immediately, to distinguish which way a current may be regarded as flowing through a conductor, and also to measure the strength of it; this can be done by means of the effects which the current produces, but the heating effect can only be used for the latter purpose, for it is found that if a current be sent first one and then the other way through a given conductor, just the same amount of heat is generated by it per second, i. e. the heating effect is *irreversible*. Further, if the current be measured in the manner described below, by means of its chemical or magnetic effect, it will be

found that the *heat generated per second is proportional to the square of the current strength.*

A current may be measured by means of the heat it produces. Any instrument for measuring current is called an ammeter (from ampère, the name of the unit of current, *vid. inf.*), and there are in use 'hot wire ammeters' whose construction depends on this principle. These consist of a fine wire, usually stretched horizontally, the middle point of which is attached to a pointer; when a current is passed through the wire, it becomes hot and sags; this causes the pointer to move round a dial, and registers the strength of a current. Such instruments are only adapted for strong currents, and, from what was said above, it follows that they are not capable of showing which way the current is flowing.

(ii) An electric current may easily be shown to produce magnetic effects in its neighbourhood. Thus if an insulated copper wire be wrapped spirally round a bar of soft iron, and a current be caused to flow through the wire, the iron will become magnetic. And again if a magnet, pivoted so as to be free to move in a horizontal plane (e. g. a compass needle), be held near a wire through which current is passing, it will in general be turned out of its normal N. and S. direction. In distinction from the thermal effects it is to be noted that the magnetic effect of a current is produced *near to, but outside* the conductor; and that it is *reversible*, i. e. when the current is made to flow the other way, the magnetic effect, whatever it is, is inverted; in the case of magnetizing an iron bar, what was the north pole becomes the south; in the case of deflecting a compass needle the deflection will take place in the opposite sense. The magnetic effect is made use of for the definition and measurement of the strength of the current. The actual definition of unit current must be deferred till later, since it involves magnetic quantities which have not yet been considered; but we may at least state the qualitative rule, since that is of use as a means of detecting the sense in which a current flows. If there be a single straight wire carrying a current, then a magnet placed near to it tends to set itself in the direction of the tangent to a circle drawn round the wire and passing through the centre of the magnet. Further, if one imagines oneself swimming in the wire, *in the direction of the current, and facing the magnet, then the north pole of*



*the magnet will be deflected towards the left.* This is known as Ampère's rule. The unit of electric current in practical use is called the *ampère*, from the name of the great French physicist just mentioned.

The magnetic effect is that most conveniently employed to measure currents by, and instruments for the purpose are called galvanometers. These instruments take many forms, but the principle involved is usually to place a small pivoted or suspended magnet at the centre of a coil of wire; the coil being arranged first parallel to the natural direction of the magnet needle, when a current is passed through it, the needle is turned out of its position to right or left according to the direction of the current, and to an extent that measures the amount of it.

(iii) An electric current produces a chemical action only when it is passed through a conductor belonging to the group already referred to as electrolytes. These consist, mainly, of aqueous solutions, and the peculiar chemical action produced may best be considered by means of a concrete example—say a strong solution of hydrochloric acid (HCl). Hydrochloric acid gas, on solution in water, suffers, like other electrolytic substances, a dissociation into parts called *ions*. These consist of either atoms or groups of atoms, associated with an electrical charge, and are consequently of two kinds, according as their charge is positive or negative; the former being called *cations*, the latter *anions*. In the case of hydrochloric acid, the dissociation is as simple as possible, since there are only two atoms in the molecule of that compound; of these the hydrogen takes the positive charge and the chlorine the negative. The two charges are equal in amount, so that the water containing hydrogen ions and chlorine ions possesses no electric charge as a whole. The last statement is always true, consequently the total (positive) charge on the cations in any electrolyte must be equal to the total (negative) charge on the anions. Moreover it is found that a very simple relation holds between the charges on the various ions: for these are always either equal to the charge on a hydrogen atom, or a whole number of times that quantity; the number in question in fact expresses the valency of the atom or group. Thus, the cation is usually a metal: H, K, Na, Tl, Ag carry one unit charge each; Ca, Ba, Zn, two each; Cu, Hg may carry one or two; Fe, either two or three, and so on. As a conve-



nient notation the positive electric charge may be indicated by a dot after the symbol, so that  $\text{Fe}^{\cdot\cdot}$  represents ferrous iron,  $\text{Fe}^{\cdot\cdot\cdot}$  ferric iron. Again, the anion usually forms all the rest of the molecule, so that some of the most common are the hydroxyl-ion  $\text{OH}$ , the chlorion  $\text{Cl}$ , nitron  $\text{NO}_3$ , acetyl-ion  $\text{CH}_3\text{COO}$ , all with a single charge; sulphion  $\text{SO}_4$  with a double charge. The negative charge is indicated by a dash, e. g.  $\text{OH}'$ .

The ions in a solution are so far free from each other's influence that when an electric force is applied to the solution, it drags them in opposite directions. If, therefore, we pass a current through the supposed strong hydrochloric acid, we shall separate the components, carrying the *cations with the current*, the anions against it. The current is led into and out of the liquid by means of metallic conductors—say small sheets of platinum foil, to which wires have been welded. These are called *electrodes*, and in particular that by which the current is led in, and consequently to which the anions are attracted, is called the *anode*, the other the *cathode*. The current in the liquid is actually constituted by the movement of the charged particles; negative charges moving in the negative direction count as well as positive charges moving in the positive direction, and produce identical effects.

The ions are led by the electric force as far as the electrodes, but there their progress is evidently stopped; and unless the current is to be stopped too, they must give up their electric charges to the electrodes, to flow round the metallic circuit. In doing so the ions become changed into ordinary matter, and will appear as such: in the example chosen the chlorions at the anode become ordinary gaseous chlorine, which is given off in bubbles; similarly the hydrogen ions give up their charge to the cathode, and become gaseous hydrogen, which is evolved from the solution. Thus happens the remarkable fact which distinguishes an electrochemical reaction from an ordinary chemical one, that the *products of decomposition appear only at the electrodes*, and separately; whereas in ordinary chemical cases they are produced together, and indifferently, in any part of the reacting substance.

Since the charge of electricity conveyed by an atom of any univalent substance is the same, that conveyed by the atomic weight taken in grams must be; and, further, in the case of multivalent substances, the equivalent weight in grams will

convey the same charge. This, then, is a fundamental quantity of great importance with regard to the chemical action of the current. The experimental laws that have led to the conclusions just mentioned as to atomic charges were first investigated by Faraday, and his results may be expressed as follows:—

(a) The amount of chemical action produced is proportional to the amount of electricity flowing, and is independent of the rate of flow (i. e. the current). Thus a current of one ampère flowing for five seconds will convey as much electricity as one of five ampères flowing for one second, and therefore liberate the same amount of any chemical product. The unit in which quantity of electricity is measured is that conveyed by one ampère in one second, and is called a *coulomb*.

(b) The amounts of various substances liberated by the same quantity of electricity are proportional to their chemical equivalents. Thus if the same current be led in turn through solutions of sulphuric acid, copper sulphate, and silver nitrate, hydrogen, copper, and silver will be the cations in these three liquids, and the weights liberated at the electrodes will be in the ratio 1 : 31.8 : 108, those being the equivalent weights of the elements in question.

Both results are included, along with the numerical observation, in the statement—

*One gram equivalent of any ion conveys 96,610 coulombs of electricity.*

Chemical action accordingly gives a means of measuring quantities of electricity, and so, indirectly, current. An instrument for this purpose is called a *voltameter*: those in most frequent use depend on the production of hydrogen, copper, or silver. The latter is the most accurate, but involves some trouble in use. The electrolysis is usually carried out in a platinum bowl which serves at the same time as cathode: in this is placed a fairly strong solution of silver nitrate; an anode of pure silver wire is arranged to dip into the liquid, and should be wrapped up in filter paper, so that no disintegrated fragments of it that may be formed during the electrolysis drop into the dish. With an 8 cm. dish and 30 per cent. solution currents up to 2 ampères may be used: the silver is deposited in a coherent crystalline form on the platinum. It must then be thoroughly washed and dried, and then weighed. Afterwards the silver is dissolved off



with nitric acid and the bowl is ready for use again. 1 coulomb deposits 0.0011172 gm. of silver.

For copper electrolysis a solution of the cupric sulphate ( $\text{CuSO}_4$ ) containing about 100 gms. of the crystalline salt per litre is the best. A platinum dish is unnecessary, as the anode and cathode may both be made of copper, and fixed vertically side by side in the solution. About 100 sq. cms. of depositing surface should be used per ampère; the weight of copper deposited on the cathode for each coulomb is 0.0003292 gm.

The water voltameter is most commonly used in Hofmann's form, as shown in Fig. 58. Dilute sulphuric acid, or dilute caustic soda, is used as electrolyte, and is decomposed between the platinum plates AA; oxygen is formed at the anode, hydrogen at the cathode, and these are collected separately in the graduated tubes BB. The volume of hydrogen should be double that of oxygen, but owing to solubility of the oxygen, and also to formation of ozone, the volume of oxygen is usually somewhat less, so that it is best to calculate the current from the hydrogen alone. The volume of hydrogen must be corrected for temperature, pressure, and presence of aqueous vapour, as described in Chap. III. When reduced in that way to  $0^\circ$  and 760 mm., the volume of hydrogen per coulomb is 0.1160 c.c. The hydrogen voltameter is particularly useful for measuring small quantities of electricity.

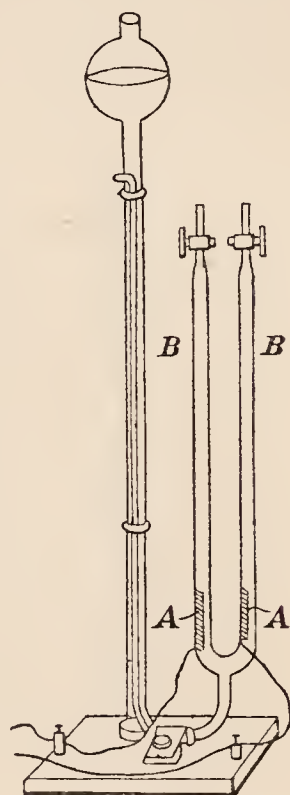


FIG. 58.

## § 2. Electromotive force.

Turning our attention now to the causes that make a current of electricity flow in a conducting circuit, we find that the phenomena cannot be described with the aid only of the conception 'quantity of electricity,' or the rate at which that flows, which is the 'current.' For in order to make a current flow it is necessary to provide a source of energy, either mechanical, as in a dynamo driven by a steam engine; or chemical, as in a voltaic cell; or thermal, as in a thermopile; and in any case the current



produced is not in itself a measure of the rate at which the supply of energy is used up—a second quantity, the *electromotive force* (E. M. F.) or *voltage*, must be measured as well. We may form a preliminary notion of this quantity, as measuring the tendency to produce current; a voltaic cell, e.g., is an arrangement always ready to produce a current, but the amount produced at any moment depends on the nature of the circuit provided for it to flow round; the electromotive force, i.e. the strength of the tendency, on the other hand, depends only on the cell itself. Thus electromotive force is comparable to pressure (in a water supply), while the electric current corresponds to the flow (of water) produced by it.

As a simple example of an arrangement for producing electric current we may take the Daniell cell. This

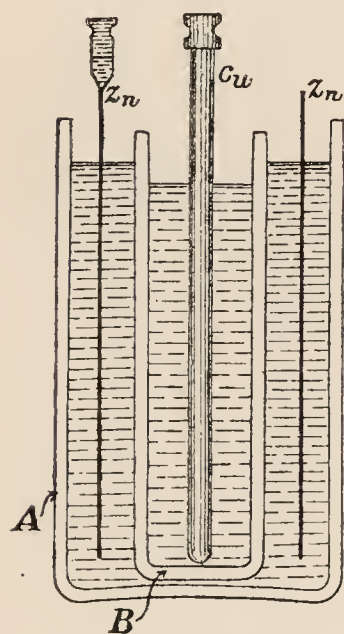


FIG. 59.

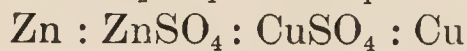
may be constructed as shown in Fig. 59.

A glass or earthenware pot A contains a sheet of zinc (Zn) bent into a cylinder for convenience; inside this stands a pot of porous earthenware B, and in it a rod of copper (Cu). The outer pot is filled up with dilute sulphuric acid (say 1 to 5 of water) or else a solution of zinc sulphate, the inner pot with a solution of copper sulphate. Wires attached to the copper and zinc plates serve to convey the current from and to the cell. The essential is merely the chemical materials used, the arrangement is otherwise a matter of convenience. We may describe the cell,

then, in its essential features as follows—



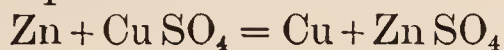
or



according to the way it is made up; the chemical action is practically the same in the two cases.

We have then two metals serving as electrodes, one (zinc) with a very strong tendency to pass into the state of ions, i.e. to acquire a (positive) electric charge and dissolve in the liquid, the other (copper) with a much weaker tendency to do so. If then a path be provided by joining Cu to Zn outside the cell by a wire a current will flow; it is carried into the liquid by zinc

which dissolves—the zinc electrode is therefore the anode—and out of it by an equivalent amount of copper which is deposited on the copper plate (consequently the cathode) being driven out by the stronger tendency of the zinc to go in. Hence the actual chemical effect is replacement of copper by zinc, and may be represented by the equation—



Now the energy liberated in such a reaction can easily be measured, for if the reaction be conducted in an ordinary chemical way all that energy is converted into heat. It has been measured and found to amount to 50,110 calories = 209,910 joules (for the quantities represented in the equation, i. e. 65 grams of zinc). This then is the store of energy available for producing a current; but we saw in the preceding section that, according to Faraday's law, 65 grams of zinc (2 equivalents) carry with them  $2 \times 96610$  coulombs of electricity. Hence the energy available is  $209910 \div 193220 = 1.087$  joules per coulomb. We have then arrived at a quantitatively exact method of expressing the tendency of the Daniell cell to produce current, i. e. of its electromotive force; it may be put in the form of a definition as follows:—

‘The electromotive force (of a cell, dynamo, &c.) is the electrical energy supplied by it per unit quantity of electricity flowing from it.’

The unit of E. M. F. adopted in practice is the joule  $\div$  coulomb, and is called a *volt*. Hence the Daniell cell has an E. M. F. of 1.087 volts. It should be noted that in the above reasoning we have assumed that all the chemical energy spent by the materials of the cell becomes electrical energy; this is nearly the case in the Daniell cell, though in many other cases it is not so. The definition of electromotive force, however, remains unaffected; it is in any case calculable if the *electrical* energy available is known.

It is sometimes more convenient to take instead of the energy the rate at which energy is spent (i. e. the power), and instead of the quantity of electricity the rate at which electricity flows (i. e. the current), and say that the E. M. F. is measured by dividing the power by the current, or in practical units that the volt = watt  $\div$  ampère.

Proceeding now to describe the commonest forms of voltaic cell, we may note first, that the essential constitution of all of them is the contact of *two different metals with an electrolyte* (or



sometimes two electrolytes). The original cell, as constructed by Volta, consisted of zinc and silver plates dipped into dilute sulphuric acid; the scheme being, therefore,  $\text{Zn} : \text{H}_2\text{SO}_4 : \text{Ag}$ . This cell is not satisfactory in its working, and is not now used, but a consideration of its defects will serve to explain the other cells that have been adapted from it. When the Volta cell is freshly made up it has an E. M. F. of about two volts. But even before any circuit is made with it, it will be found that zinc dissolves in the acid—at least if ordinary (impure) zinc be used. This is due to what is called *local action*: small particles of other metals present in the zinc constitute, with that metal, little local voltaic cells all over the plate. Now to get current out of a cell it is necessary to connect the electrodes by a conductor: if this conductor is very short and thick the greatest possible current will be obtained—the cell is said to be *short-circuited*: that would be the case with the Daniell cell described above, if Cu and Zn were connected by a thick wire; still better if the two electrodes touched one another. But when an impure zinc plate is put into acid, the little local cells formed are all short-circuited, since all their electrodes form part of one metal plate, and accordingly electrolytic action will take place, as may be seen by the bubbles of hydrogen given off all over the plate. The current so produced will of course all be wasted in the zinc plate and serve no useful purpose. If a plate of pure redistilled zinc be used the local action will not take place; and the same thing may be accomplished more economically in practice by cleaning the zinc with acid and rubbing it over well with mercury; by this means the surface of the plate is covered by a uniform layer of zinc amalgam, and the local differences of composition which cause local action avoided. Hence voltaic cells in general should be made with amalgamated zinc plates, unless chemically pure zinc is used.

Next it will be found that if the zinc and silver plates be connected so as to cause a current to flow, this current will rapidly fall off in strength, and if the voltage of the cell be tested after a few minutes' working it will be found much less than two—the cell is said to be *polarized*. The cause lies in the production of hydrogen on the cathode—the silver plate. Current in an electrolyte necessarily consists in the movement either of positively charged ions in the direction of the current, or negative ones in



the contrary direction: so in this case the current is conveyed into the electrolyte by the zinc which dissolves; but to convey it out at the cathode there is no negative ion to go into solution, for the silver plate contains no materials for such, consequently a positive ion must go out of solution, and the only kind that is to hand is the hydrogen of the acid. Accordingly hydrogen will be liberated on the cathode, as in a water voltameter.

The effect of the hydrogen is twofold. In the first place, gases are all non-conductors of electricity, so that wherever a bubble of hydrogen appears the conducting circuit will be interrupted, and if much hydrogen is produced the channel for the current to flow through will be much obstructed, and the flow of current correspondingly diminished: in the language of the next section we may say that the *internal resistance* of the cell is *increased*. But even before any visible amount of hydrogen appears on the plate, its chemical influence becomes apparent; the hydrogen may in the first instance be regarded as dissolving in the silver—this phenomenon is well known in the case of platinum and palladium, the latter of which will absorb many times its own volume of hydrogen, but even in the case of silver a minute amount is probably absorbed. Now the dissolved hydrogen practically converts the silver into a hydrogen plate, and alters the character of the cell altogether; this is the phenomenon known as *polarization*. The dissolved hydrogen has a considerable tendency to go back into the ionic form, and indeed hydrogen combining with the elements sulphur and oxygen to form  $\text{H}_2\text{SO}_4$  would evolve not much less heat than zinc in forming the corresponding compound  $\text{ZnSO}_4$ ; so that instead of the energy of combination of the zinc being available for producing electromotive force, as at the moment when the cell first comes into action, the energy of recombination of the hydrogen must be deducted, and the E. M. F. corresponds merely to the difference in the tendencies of zinc and hydrogen to ionize: instead of two volts it is a fraction of a volt.

Thus in order to construct a workable cell it is necessary to avoid polarization. This can be accomplished by mechanical means, such as brushing away the hydrogen, but only very imperfectly; it can be better done by chemical action, oxidizing away the hydrogen as it is formed. Nitric acid, potassium

bichromate, and manganese dioxide are the oxidizing agents mostly used. The cells depending on their action are—

*Grove's.*      $\text{Zn} : \text{H}_2\text{SO}_4 \text{ (dilute)} : \text{HNO}_3 \text{ (strong)} : \text{Pt}$

The sulphuric acid may be about 1 to 5 or 1 to 7 of water, the nitric acid not diluted, in order to get the highest electromotive force. The two liquids must be separated, most conveniently by a jar of porous earthenware; for if the nitric acid came in contact with the zinc it would dissolve it even when the circuit was not completed. Silver cannot be used for cathode as it is attacked by nitric acid; platinum is therefore used instead. The Grove cell has an E. M. F. of about 1.9 volts, is capable of giving fairly strong currents, and is steady in working for two or three hours; but it does not last long, and must be taken to pieces after using; moreover it gives off unpleasant nitrous fumes.

*Bunsen's.*      $\text{Zn} : \text{H}_2\text{SO}_4 \text{ (dilute)} : \text{HNO}_3 \text{ (strong)} : \text{C}$

Identical with Grove's, except that a stick of hard carbon is substituted for the expensive platinum; its efficiency is practically the same.

*The bichromate.*      $\text{Zn} : \text{H}_2\text{SO}_4 + \text{K}_2\text{Cr}_2\text{O}_7 \text{ (dilute)} : \text{C}$

The solution may consist of 200 grams of sulphuric acid and 80 grams of potassium bichromate, made up to a litre of water. No porous pot is used, because the solution does not act much on the zinc when the cell is not in use; but it is desirable to have an arrangement for drawing the zinc up out of the liquid when the cell is done with, to prevent waste. Its electromotive force is from 1.8 to 2 volts, and is not quite so steady as that of a Grove; the cell is used for similar purposes.

*Leclanché's.*      $\text{Zn} : \text{NH}_4\text{Cl solution} : \text{MnO}_2 : \text{C}$

The solution not being an acid does not attack the zinc at all when the circuit is not completed. Hence the cell may be left fitted up for any length of time, until the solution needs renewing. It is not suited for giving strong currents, as it polarizes a good deal; the manganese dioxide, used as depolarizer, is a solid, and its action consequently slow, but if the cell is only used for a short time and then left to itself, the polarization is removed, and the E. M. F. recovers to its original value; hence it is particularly useful for electric bells. The Leclanché may be conveniently used for most electrical measurements: its E. M. F. is about 1.4 volts.



*Dry cells* are Leclanché's made up with sawdust or some other substance to prevent the liquid getting spilt, and sealed up. They are very convenient where small currents are required, as they need no attention, produce no fumes, and give no opportunity for mess.

*Daniell's.* This has already been described. In it polarization is not so much counteracted as avoided altogether. It gives stronger currents if made up with sulphuric acid, but is steadier in action if made with zinc sulphate solution. In the latter case hydrogen ions have practically nothing to do with its action, for the current is carried into the solution at the anode by zinc ions and out at the cathode by copper ions. It is necessary to keep the copper sulphate solution away from the zinc to prevent direct deposition of copper on it; this can be done by a porous pot as above described, or by making the zinc sulphate solution much denser than the copper sulphate and pouring the latter carefully on top of the former. The cell always breaks down in the end by diffusion, and so needs to be set up afresh from time to time.

*Clark's standard cell.*  $\text{Zn} : \text{ZnSO}_4 \text{ (saturated)} : \text{Hg}_2\text{SO}_4 : \text{Hg}$

The constitution of this is very similar to that of the Daniell cell, but the depolarizer, the mercurous sulphate, is an almost insoluble substance. This has as one result that the cell polarizes if at all strong currents are taken from it, and requires some time to depolarize; but it is not intended for giving appreciable currents, but only as a standard of electromotive force; while the insolubility of the mercurous sulphate keeps it from action on the zinc. The Clark cell, consequently, if made up with sufficient care, will keep indefinitely, and forms a perfectly reliable standard of E. M. F. The practical construction is shown in Fig. 60. A short wide test tube contains the mercury *m* for cathode; connexion with this is made by means of a platinum wire *p* sealed through the glass. Over the mercury is poured a neutral saturated solution of zinc sulphate, to which mercurous sulphate has been added to form a thick paste. This should be so full of crystals as to set firmly and so prevent any chance of

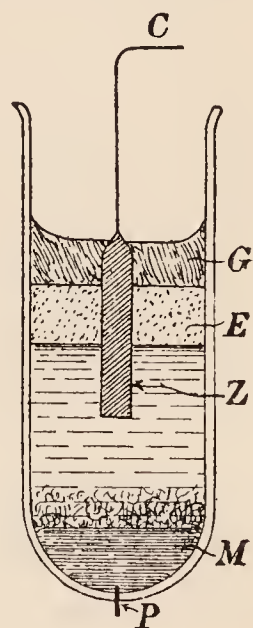


FIG. 60.



spilling the mercury if the cell be upset ; the anode of pure zinc *z* is held in the solution by a cork *E*, a copper wire *c* soldered on to it serving to make connexion ; and the whole is closed air-tight by marine glue *G*. The E. M. F. is 1.433 volts at 15°, and falls off by 0.0012 volt per 1° rise of temperature.

The above are so-called primary cells, i. e. arrangements for producing electric current by the expenditure of chemical materials : when the zinc, or acid, or other material in them is used up, it must be replaced by fresh. There are also, however, *secondary cells* or *accumulators*, in which the chemical materials after use are reproduced by running a current in the opposite sense through the cell. The most important of these is the ordinary lead accumulator. Its action, in outline, is as follows : When the cell is fully charged, the anode plate is of lead, the cathode of lead peroxide (held together by means of a grid of metallic lead), the electrolyte dilute sulphuric acid (about 1 to 5 of water), i. e.  $\text{Pb} : \text{H}_2\text{SO}_4 : \text{PbO}_2$ . The action of the cell is to oxidize the anode and reduce the cathode, so that both tend to become  $\text{PbO}$ , or rather, in presence of sulphuric acid,  $\text{PbSO}_4$  ; when the two plates are alike the E. M. F. of course vanishes ; but by running current from an outside source in the opposite sense,  $\text{Pb}$  and  $\text{PbO}_2$  are reformed—this is known as charging the cell—and it is then ready for a further discharge. The accumulator possesses the advantages of all the other cells combined ; it is capable of providing very strong currents, its E. M. F. remains for a long time constant ; it does not waste when not in use, and is consequently always ready for use, and does not require much attention. The disadvantage is of course the necessity for charging ; this is usually done by means of a dynamo. When continuous current from a public supply is available, it may be used very conveniently, and the cells should be charged once a week, or thereabouts. Their E. M. F. is normally 2.05 volts ; during charging 20 per cent. more is required, and the cells should not be discharged any more when their voltage has dropped some 10 per cent. below the normal amount.

### § 3. Resistance.

It appeared in the preceding section that the electromotive force in a circuit, being located where there is a source of electrical energy (cell, dynamo, &c.), could be calculated by means of

the amount of energy available ; but the strength of the current which flows through the circuit depends not only on the electromotive force in it, but also on the nature of the circuit itself. The latter influence may be described by means of the *resistance* offered by the various conductors to the flow. Thus if in a given circuit it is found that  $x$  volts are required to make one ampère flow,  $x$  may be regarded as a measure of the resistance offered to the flow ; for if in some other circuit it was found that  $2x$  volts were needed to maintain the same current, we should naturally say that the second circuit offered twice as much resistance. Now it is found that, confining attention to a single circuit, the amount of current produced in it is proportional to the electromotive force applied. This most important result, which is known as *Ohm's law*, is verified by experiment quite strictly, both in the case of metallic and electrolytic conductors ; at the same time it must not be supposed that it is an *a priori* law involved in the nature of current and electromotive force ; that that is not so may be clearly seen by considering the analogous case of the flow of water through a pipe—the rate of flow is here not proportional to the pressure driving it. The law is, as a matter of fact, found to be true not only for complete circuits, but for every separate conductor and material. So it may be put in the form—

*'The current produced in any conductor is proportional to the E. M. F. applied to it.'*

In accordance with this law, therefore, the resistance of the conductor may be defined as the (constant) ratio of E. M. F. applied to current produced : or in symbols—

$$R = \frac{E}{C}$$

where  $E$  = electromotive force,  $C$  = current,  $R$  = resistance. The *resistance* of a wire is therefore to be *measured in volts per ampère* ; this unit is called the *ohm*.

Sometimes it is more convenient to use, instead of resistance, its reciprocal, i. e. the ratio of current to E. M. F. This is called the *conductance*, and is *measured in ampères per volt*, or *mhos*. For example, a sixteen-candle power electric lamp adapted for use with a public supply of electricity at 110 volts may take 0.5 ampère. Accordingly its resistance is  $110 \div 0.5 = 220$  ohms, or its conductance  $\frac{1}{220}$  mho.



The resistance of an ordinary cylindrical conductor, or wire, is found experimentally to be proportional to the length, inversely to the area of cross section, and to depend very much on the nature of the material it is made of. The latter influence may be described by taking as standard a length of 1 cm. of a conductor whose cross section is 1 sq. cm.; the resistance of this is called the *resistivity* (or specific resistance) of the material. Hence we may put—

$$R = \frac{\rho l}{a}$$

Here  $\rho$  = resistivity,  $l$  = length,  $a$  = area of cross section; e. g. if it be known that the resistivity of pure copper is  $1.6 \times 10^{-6}$  ohms, it is desired to find the resistance of a No. 24 copper wire 20 metres long. Here  $l = 2,000$  cms.,  $a = 0.00245$  sq. cms., according to the wire gauge table, so that  $R = 1.6 \times 10^{-6} \times 2000 \div 0.00245 = 1.36$  ohm.

Similarly the conductance of a specimen 1 cm. long and 1 sq. cm. in cross section is called the conductivity, so that the conductance of a wire (or tube of electrolyte) = conductivity  $\times$  cross section  $\div$  length.

The resistivity of all substances depends on temperature. Accordingly in the following table the temperature coefficient is given; this means the *increase of resistance per 1° expressed as a fraction of the resistance at 0°*.

	Resistivity at 0° (ohms).	Temperature coefficient.
Silver . . . . .	$1.47 \times 10^{-6}$	+ 0.00400
Copper . . . . .	$1.56 \times 10^{-6}$	0.00428
Aluminium . . . . .	$2.66 \times 10^{-6}$	0.00435
Platinum . . . . .	$10.92 \times 10^{-6}$	0.00367
Iron . . . . .	$9.06 \times 10^{-6}$	0.00625
German silver alloy . . . . .	$20. \times 10^{-6}$	0.00044
Platinum silver alloy . . . . .	$24. \times 10^{-6}$	0.00031
Manganin alloy . . . . .	$42. \times 10^{-6}$	negligible
Mercury . . . . .	$94.07 \times 10^{-6}$	0.00088
Graphite and electric light carbons	$2400 \times 10^{-6}$ and upwards	- 0.0005

It will be observed that all the metals increase in resistance with rise of temperature, and further that the various pure metals have mostly temperature coefficients of about  $0.00366 = \frac{1}{273}$ , which is the coefficient of expansion of a gas. That is, the resistance of a pure metal varies in nearly the same way as the volume of a gas when the temperature is altered, and accordingly



would vanish, or at least become very small at the absolute zero. Carbon, on the other hand, diminishes in resistance when the temperature rises, so that the resistance of an incandescent lamp is much less when hot than when cold.

Electrolytes have all much greater resistance than metals ; one of the best conducting is strong nitric acid, but the resistivity even of this is of the order of 1 ohm, i. e. roughly a million times greater than copper. As a general rule the conductivity of solutions goes hand in hand with the amount of dissolved substance they contain, so that the weaker a solution is made the greater its resistance is. 0.6% salt solution, which is commonly taken to correspond in salt content with the blood, has a resistivity of about 80 ohms. Most of the tissues have probably a higher resistance than this.

Absolutely pure water is almost a non-conductor, having a resistivity of about 25,000,000 ohms. Almost all the electrolytes conduct better hot than cold, the rate of change being very rapid, usually 0.02 per degree.

Even the value for water is however far exceeded by the actual insulators—glass, silk, gutta-percha, and so on—e. g. the resistivity of mica has been estimated at  $10^{14}$  ohms. The difference between this value and those for metals is so enormous that a No. 24 copper wire stretching all round the earth would have less resistance than a thin mica wad, so that the current would rather travel by that route, 40,000 kilometres long, than from one face of the mica to the other. It is this fact that renders electrical energy so superior to any other kind in practical availability, for the current conveying the energy can be guided by wires to any point desired, however distant, without excessive leakage ; it is thus possible to drive an electromotor by means of current generated a hundred kilometres away ; and the small currents needed for telegraphic signalling can even be carried in practice across the ocean.

It is easy to construct permanent standards of electrical resistance, since the resistance of a wire (at least when properly annealed) remains quite constant but for temperature variations. To avoid the influence of these last, such standards are made of some alloy whose temperature coefficient is small, usually nowadays of manganin. Coils of resistance equal to 1, 2, 3, 10, 100, 1,000, or any desired number of ohms are made ; and usually these

are made up into *resistance boxes*, arranged as shown in Fig. 61 (*a* plan, *b* elevation). Here P Q R S T are brass blocks, to which are soldered the ends of a coil, a well-fitting conical brass plug fills up the space between the blocks, and the current flows through this with no appreciable resistance; but when the plug is removed the current can only flow through the coil, and so much resistance is thrown into the circuit. The coils are made in sets which can be combined so as to give any desired total, like a set of weights.

For very large resistances it is more convenient to use either carbon or an electrolyte. Megohm (1,000,000 ohms) standards are made consisting of a streak of graphite (i. e. a lead-pencil mark) on an ebonite plate. Liquid resistances of 1,000 ohms or more

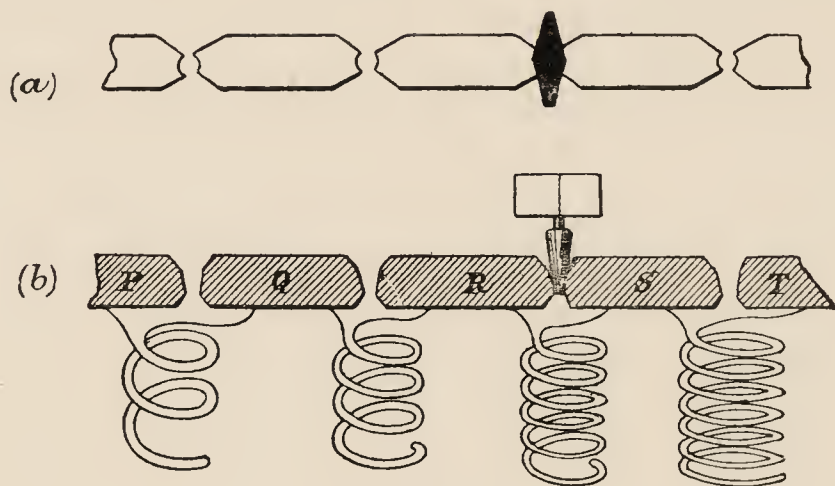


FIG. 61.

can easily be constructed by filling a glass tube with zinc sulphate solution, and fitting it with electrodes of pure amalgamated zinc, since such electrodes do not polarize appreciably; a liquid resistance would however be inconstant if it polarized, for the reasons discussed in the last section, unless it were used with an alternating current (i. e. one that oscillates to and fro instead of flowing always in the same direction).

Besides standards of resistance for purposes of measurement it is desirable to have variable resistances, by which to adjust the strength of the current in a circuit. For weak currents ( $\frac{1}{10}$  ampère and less) resistance boxes may be used, but it is simpler to have a wire of variable length, such as that of the apparatus (rheostat)



shown in Fig. 62. Here the wire is wound spirally on a drum, the ends being connected to the binding screws A, B. The third screw C is connected to a brass spring D, which slides along a bar. According to the position of the sliding contact, more or less of the wire is interposed electrically between A and C.

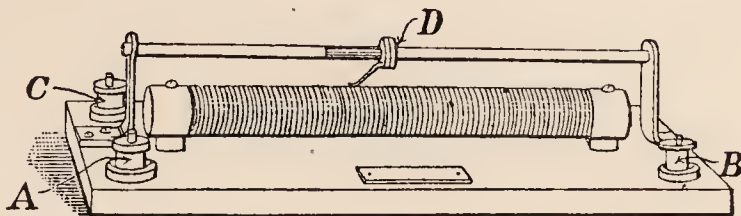


FIG. 62.

Another piece of apparatus for this purpose is the Varley rheostat (Fig. 62 a), consisting of a pile of loose carbon discs; when these are pressed closer together by a screw their resistance diminishes. For very small currents, such as those occurring in nerve-muscle work, it is convenient to use liquid resistances made as described above, but with one of the electrodes fixed on a rod so that it can be pushed in and out of the tube. The longer the distance between the electrodes is made the greater, of course, is the resistance offered by the liquid.

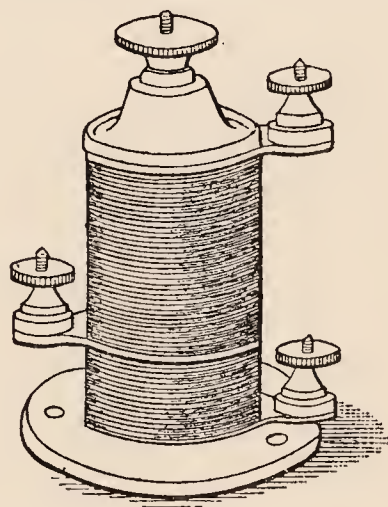


FIG. 62 a.

An actual electric circuit is made up of various conductors—battery, connecting wires, galvanometer, resistance coils, and what not, and in the simplest case the current flows through all these in turn. Since Ohm's law is applicable to each of them, it is applicable to the circuit as a whole, and to get the current flowing we have only to divide the electromotive force by the total resistance in the circuit. Further, in accordance with Ohm's law the electromotive force is spent partly on each conductor, and in proportion to the resistance overcome. This may be elucidated by an actual example. Thus suppose (Fig. 63) an accumulator BB' of 2.05 volts, joined up to a small lamp L of 3.6 ohms resistance, an ammeter A of 0.3 ohm, and let the internal resistance of the cell (i. e. the resistance from terminal to terminal, mainly made up of that of the electrolyte between the plates) be 0.2 ohm, that of the conducting wires negligible. The



total resistance is  $3.6 + 0.3 + 0.2 = 4.1$ , hence the current is  $2.05 \div 4.1 = 0.5$  ampère. If now a voltmeter be attached to the ends of the lamp, it will indicate that part of the E. M. F. used to drive the current through the lamp alone, viz.  $0.5 \text{ ampère} \times 3.6 \text{ ohms} = 1.8 \text{ volts}$ ; if it be attached to the ends of the ammeter it

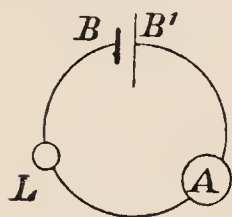


FIG. 63.

will read, similarly,  $0.5 \times 0.3 = 0.15$  volt. If however it be attached to the terminals of the cell it will read 1.95 volts, for the current in flowing through the cell would require  $0.5 \times 0.2 = 0.1$  volt to drive it, and on that account the terminal  $B'$  (the lead plate) where the current enters would be 0.1 volt above  $B$  (the peroxide plate); but the cell is itself the source of electromotive force, so that in passing from the lead to the peroxide plate there would on account of chemical action be a rise of 2.05 volts; hence on the whole  $B$  is higher than  $B'$  by 1.95 volts, and  $B$  is the positive terminal. Of course if no current were flowing from the cell a voltmeter attached to it would read 2.05 volts.

The point here involved is very clearly brought out by comparing a voltaic cell (Cu Zn, Fig. 64) with an electrolytic cell (Cu Cu, Fig. 65). When currents flow through these there is in

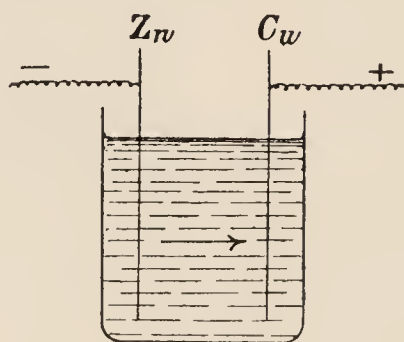


FIG. 64.

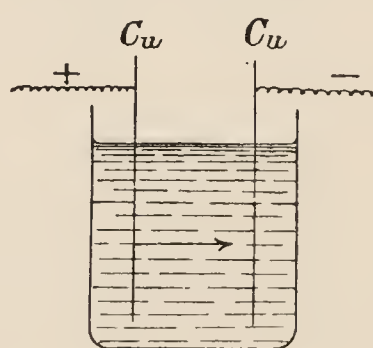


FIG. 65.

each a drop in volts in passing from the point where the current enters (anode) to that at which it leaves (cathode) on account of resistance overcome; in the electrolytic cell there is nothing to compensate this, so the anode is +<sup>ve</sup>. But in the voltaic cell there is an electromotive force acting in the direction in which the current is flowing (from Zn to Cu through the cell), which causes the cathode (Cu) to be the +<sup>ve</sup>.

To drive a current through an electrolytic cell needs a certain

voltage on account of the resistance of the solution, and when there is no chemical work to be done that is all; i. e. when copper sulphate is electrolyzed between copper plates, just as much copper is dissolved at one pole as is deposited at the other, so that no chemical work is done. But when that condition does not hold a further voltage is required. Whenever polarization occurs this is the case, e. g. when dilute sulphuric acid is decomposed between platinum electrodes; here work is done in decomposing the solution and generating oxygen and hydrogen, and it is found that about 1.7 volts has to be spent to do this, in addition to anything required to merely overcome the resistance of this electrolytic cell. The amount thus needed, which may be called the *back electromotive force of polarization*, can be approximately determined, like the E. M. F. of a voltaic cell, by dividing the chemical energy involved in the reaction by the quantity of electricity which, according to Faraday's laws, is conveyed: only here the chemical work is done by the current, whereas in the voltaic cell the current is produced at the expense of chemical energy. An electrolytic cell is therefore, in a sense, the converse of a voltaic cell.

Indeed the same piece of apparatus may serve for both. Thus a lead accumulator when charged is a voltaic cell, and is capable of giving current at the expense of its store of chemical energy; but when run down it has to be treated like an electrolytic cell, and a current run against its own E. M. F. so as to 'charge' it again, i. e. restore the chemical energy which it has lost. Here the peroxide plate is in both cases positive to the other, but it acts as cathode during the discharge, anode during the charge.

We arrive then at the following results as to distribution of energy in electric circuits:—

(i) The total electromotive force  $E$  in a circuit is spent (*a*) in overcoming any back electromotive force  $E'$  due to polarization that may exist in the circuit, and (*b*) the remainder of it in driving a current  $C$  against the resistance  $R$  of the circuit, so that

$$E = E' + CR$$

If there is no back electromotive force  $E$  becomes  $= CR$ .

(ii) Bearing in mind that electromotive force means power  $\div$  current (p. 193) it follows that the total electrical power (or activity)  $A$  in a circuit is spent (*a*) in effecting chemical decom-

position against the electromotive force of polarization, the power spent being  $E'C$ ; (b) in producing heat in the circuit at the rate  $CR \times C$ , so that

$$A = E'C + C^2R$$

Here the power will be given in watts if the ampère, volt, and ohm are the units employed.

(iii) The work  $W$  done is of course the power multiplied by the time ( $t$ ), so that

$$W = E'Ct + C^2Rt$$

where the first term on the right-hand side is the work done in chemical decomposition, and the second the heat produced in the circuit; if the time is given in seconds both these will be given in joules. To calculate the heat in calories the number must be divided by 4.2.

It may be added that an electromotor (vid. inf. p. 238) behaves like an electrolytic cell in giving a back electromotive force. Hence in the preceding equations  $E'$  may also be taken to mean that, and  $E'C$  will be the amount of electrical power spent in doing mechanical work by means of the motor.

Example: A motor whose resistance is 6 ohms is connected to a supply of electricity at 110 volts, and it is observed that 2 ampères flow through it. Consequently the voltage used in overcoming resistance is  $(CR) 2 \times 6 = 12$ , leaving  $110 - 12 = 98$  volts for actually doing work by the motor ( $E'$ ). Hence further, the power obtained from the motor (neglecting friction and other mechanical imperfections) is  $98 \times 2 = 196$  watts; the power spent in heating is  $(C^2R) 2 \times 2 \times 6 = 24$  watts, and the total power taken from the supply ( $A$ ) is  $196 + 24 = 220$  watts. Again, the heat produced per hour (3,600 seconds) is  $24 \times 3600 = 86400$  joules or  $86400 \div 4.2 = 20570$  calories.

#### § 4. Distribution of currents.

We have so far considered only the simplest kind of circuit, one in which all the current flows in turn through all the conductors. But it often happens in practice that branches and networks of conductors exist, so that the simplest formulation of Ohm's law is not applicable. We have to consider in this case how the current in each branch may be calculated. All such cases may be reduced, so far as the resistance is concerned, to



two: the conductors are arranged either in series or in parallel, or it may be some combination of the two. *Series* means that the same current flows through each of them in turn (Fig. 66, *a*), *parallel* that the conductors offer alternative paths to the current (Fig. 66, *b*). In order to determine the joint resistance, that is the total resistance from A to B or from C to D in the figure, we have the following rules:—

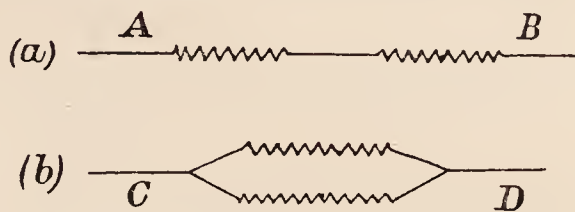


FIG. 66.

(a) 'The resistance of any number of conductors in series is the sum of their separate resistances.'

This almost obvious rule was assumed in the course of the preceding section.

(b) 'The conductance of any number of conductors in parallel is the sum of their separate conductances.'

Thus if  $p$   $q$   $r$  be the resistances of the separate conductors, their conductances are  $\frac{1}{p}$   $\frac{1}{q}$   $\frac{1}{r}$ ; the total conductance is therefore  $\frac{1}{p} + \frac{1}{q} + \frac{1}{r}$ , and the total resistance  $\frac{1}{\frac{1}{p} + \frac{1}{q} + \frac{1}{r}}$ .

One of the most familiar cases of putting conductors in parallel is in the ordinary way of wiring a house for electric lighting. This is to place each lamp independently across the mains, i. e. from the main conductor leading from the positive end of the dynamo to the other main conductor leading to the negative end. All these lamps are in parallel to one another, therefore; suppose a supply at 110 volts, and that 18 lamps each of 200 ohms, one large lamp of 50 ohms, and a heater of 20 ohms are put in. Then the conductances are, each small lamp  $\frac{1}{200} = 0.005$  mho, hence the 18 small lamps are  $18 \times 0.005 = 0.09$  mho, the large lamp 0.02 mho, the heater 0.05 mho, making a total of  $0.09 + 0.02 + 0.05 = 0.16$  mho. The total resistance when all are switched on is thus  $1 \div 0.16 = 6.66$  ohms, and the total current = E. M. F.  $\div$  resistance, or = E. M. F.  $\times$  conductance, i. e.  $110 \times 0.16 = 17.6$  ampères.

Another common case is a galvanometer and its shunt. When a galvanometer is too sensitive for the purpose in hand a wire

resistance is arranged in parallel with it, as shown in Fig. 67; then only part of the current flows through it. If  $G$  and  $s$  be the resistance of the galvanometer and shunt respectively, their combined resistance is  $\frac{I}{\frac{1}{s} + \frac{1}{G}} = \frac{SG}{s + G}$ . Usually  $G$  is very much

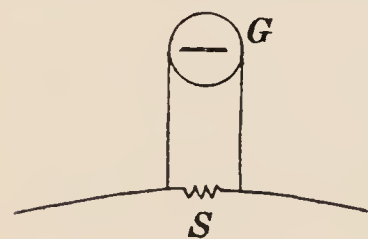


FIG. 67.

greater than  $s$ ; in that case the combined resistance is practically that of the shunt alone (i.e. the conductance of the galvanometer can be neglected by comparison with that of the shunt).

When conductors are put in parallel there is evidently the same voltage applied to each, seeing that they connect the same points (e. g.  $c$  to  $d$  in Fig. 66); hence in accordance with Ohm's law the current flowing through each must be proportional to its conductance. Thus the currents through galvanometer and shunt respectively, and the total, will be

galv. current : shunt current : total current ::  $\frac{I}{G} : \frac{I}{s} : \frac{I}{G} + \frac{I}{s}$ .

Hence the fraction of the total current flowing through the galvanometer is  $\frac{I}{G} \div \left(\frac{I}{G} + \frac{I}{s}\right) = \frac{s}{s + G}$ . Usually the shunt is made to have a resistance  $\frac{1}{9}$ , or  $\frac{1}{99}$ , or  $\frac{1}{999}$  of that of the galvanometer, i. e. 9, 99, or 999 times its conductance, and accordingly  $\frac{1}{10}$ ,  $\frac{1}{100}$ , or  $\frac{1}{1000}$  of the total current flows through the galvanometer according to the shunt used.

A 'short circuit' is a shunt of very low resistance, so that when a piece of apparatus is short-circuited almost none of the current flows through it, and it is practically cut out of use, except in the case of the battery. The current of course all comes from it in any case, and a short circuit merely makes it very large (and probably damages the battery).

It sometimes happens that the voltaic cells themselves have to be arranged in parallel or in series; then their internal resistances can be combined according to the preceding rules just like any other resistances; but the electromotive forces require separate treatment. Three arrangements may be considered. (a) Cells are placed in *series*, i. e. the zinc of the first is connected to the copper or carbon of the second, the zinc of that to the

copper or carbon of the third, and so on ; then the electromotive force is simply the sum of the separate electromotive forces. Example : one Daniell cell of 1.08 volts is found incapable of electrolyzing water ; two such are put in series, so giving 2.16 volts, and the electrolysis is found to take place (Fig. 68).

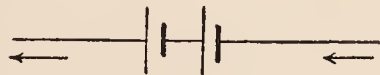


FIG. 68.

(b) Two cells are said to be in *opposition* when they are arranged as in series, except that one of them is turned the other way round. The total electromotive force is now the difference between the two (Fig. 69). Example : in order to charge a battery of 40 accumulators they are connected in series, and then put in opposition to a dynamo of 110 volts. If during charging each cell has an E. M. F. of 2.45 volts, the total E. M. F. of the battery is  $40 \times 2.45 = 98$  ; hence the total E. M. F. in the circuit is  $110 - 98 = 12$  volts. It is this that is effective in driving the current.

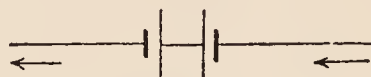


FIG. 69.

(c) Cells are arranged in *parallel* (Fig. 70) by joining together all their positive terminals and all their negative terminals, and attaching these joint electrodes to the external circuit. This should only be done when all the cells are alike, and so have the same electromotive force. In this case the combined E. M. F. is merely that of a single cell ; in fact the group practically forms one cell with bigger electrodes. The only effect, then, as compared with a single cell, is to reduce the internal resistance ; but sometimes that is the most effective way of increasing the current in a circuit.

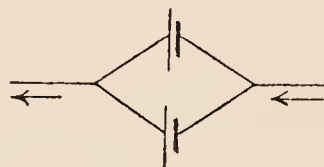


FIG. 70.

So far we have paid no special attention to the internal resistance of a cell, as it is constituted in the same way as the resistance of any other conductor, and depends on the same factors. But it plays such an important part in determining the efficiency of a battery as to deserve some separate mention.

To take a practical example, let us suppose two dry cells, each of 1.4 volts, and 3 ohms internal resistance. Either cell by itself is capable of giving as a maximum  $1.4 \div 3 = 0.47$  ampères, for if it were short-circuited, i. e. the poles connected by a wire of negligible



resistance, then the total resistance in the circuit would be the internal resistance only, and the current is found, according to Ohm's law, by dividing this into the electromotive force. If the two cells were put in series and short-circuited, the total E. M. F. would be 2.8 volts, and total resistance 6 ohms, and the current would be  $2.8 \div 6 = 0.47$ , the same as before ; but if the two were put in parallel, while the E. M. F. would remain 1.4 volts, the joint resistance would be reduced to 1.5 ohms, and the current  $1.4 \div 1.5 = 0.94$  ampère, would be doubled. Hence we see that if it is desired to get the largest current through a very small external resistance it is best to put the cells in parallel. On the other hand, if the external resistance is very large, the rule is reversed. Suppose it to be 1,000 ohms, then we may practically neglect the internal resistance by comparison with this, and we find that one cell would give  $1.4 \div 1000 = 0.0014$  ampère, and any number of cells in parallel would only give the same ; but the two cells in series would give  $2.8 \div 1000 = 0.0028$ , i. e. twice as much current. Thus, when the external resistance is very large, the cells should be put in series.

Accumulators have very low internal resistance : a good-sized one may have  $\frac{1}{100}$  ohm or less ; consequently they are capable of giving very strong currents.

The extreme contrary case is to be found in a nerve-muscle preparation ; this acts as a voltaic cell, giving an electromotive force of about 0.1 volt ; but it has a very high internal resistance—100,000 ohms possibly. Hence, any ordinary electric circuit, even a galvanometer of 10,000 ohms, acts practically as a short circuit to it, and all that has been said of short circuits is applicable ; it makes, in fact, very little difference what path is provided for the muscle-current, the amount obtained will be practically the same if any ordinary wire circuit at all is used.

### § 5. Methods of measurement.

*Current.* Of the quantities we have so far discussed, one, the intensity of an electric current, is commonly measured directly by an appropriate instrument based on one of the effects it produces ; whereas electromotive force and resistance are usually found by means of more complex and indirect processes, essentially involving some current-measuring device ; it is necessary, therefore, first to deal with current measurement.

The principles on which such instruments can be constructed were referred to in general terms in § 1, and it appears that they fall under two main heads—thermal and magnetic effects. The use of the former is, however, restricted to a few electrical engineering instruments, so that we may say the only current measures with which we have to deal belong to the class of galvanometers, using that word to cover any instrument in which magnetic effects are used for the purpose in question. To give the precise theory of any galvanometer would be anticipating the subsequent chapter on electromagnetics: it will be sufficient here to mention points of theory only so far as they are indispensable in going over the practical construction of galvanometers. Three groups of instruments may be made, according as the instrument contains (i) a fixed coil of wire through which the current flows, and in consequence tends to rotate a magnet,

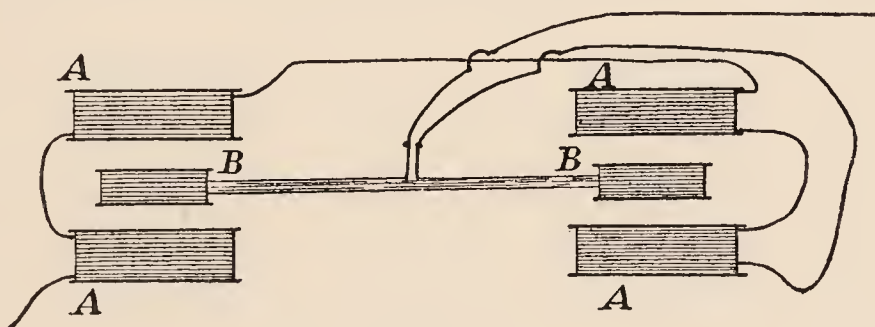


FIG. 71.

either suspended or pivoted so as to be free to move; (ii) a suspended or pivoted coil through which the current flows, and which, placed between the poles of a magnet, tends in consequence to rotate; (iii) a fixed coil and a movable coil, through both of which the current is led.

For measuring currents of intensity down to about one milli-ampère ( $\frac{1}{1000}$  ampère) direct reading instruments based on these various plans are in common use. Designed primarily for engineering purposes, such 'ammeters' usually possess an accuracy of the order of one per cent., and can be used for physical experiments except when the highest order of exactness is required; in physiological experiments always, when such large currents are to be measured. There are, moreover, instruments of a more exact construction, intended mainly for standardizing others, among which may be mentioned Lord Kelvin's 'current balances'; these (Fig. 71) consist essentially of four fixed coils AAAA, and



two movable ones BB, the current being led through all in turn. The two movable coils are attached to a frame supported at its centre so as to be free to turn like a balance. Of the two fixed coils on the right one attracts the movable coil near it, the other repels, so that both tend to raise the movable coil which lies between them; similarly, the two at the left both tend to lower the movable coil which lies between them. Hence, all four coils tend to tilt the beam in the same direction; to balance this a small weight is hung on the beam and slid along it (as in a steel-yard) till the beam lies exactly horizontal. The moment of the weight then serves to measure the current flowing. This instrument, though depending on what may by analogy be called the magnetic attractions and repulsions of electric currents, does not involve any actual steel magnets, and is consequently not exposed to errors due to accidental weakening of the magnet, like most direct reading ammeters. It must be noted that, like all other instruments in which the current passes through both a fixed and a movable coil, the scale reading is proportional to the square of the current, for doubling the current doubles the force exerted on account of the fixed and movable coils separately, making it on the whole four times as great. In such an instrument the motion is always one way, whichever way the current flows, but that is for some purposes an advantage, because it allows of measuring an alternating current, i. e. not a steady flow of electricity in one direction, but an oscillation of electricity to and fro along a conductor.

Amongst instruments for measuring comparatively large currents may be mentioned, on account of its theoretical simplicity, the tangent galvanometer. This consists of a circular coil of wire fixed on a frame. At the centre of the coil, which is commonly made from 10 to 20 cms. in diameter, is fixed a small horizontal graduated circle; here a short compass needle, provided with a long, light pointer of aluminium or glass, is pivoted. The needle being free to move in a horizontal plane sets north and south; but when a current is led through the coil there is a tendency for the needle to set itself at right angles to the coil. Accordingly, to use the instrument the coil is rotated till it is parallel to the needle (i. e. in the magnetic north and south line); then the current exerts a couple on the needle, tending to make it lie east and west, and this, together with the couple due to the



earth's magnetic action, will cause the needle to rest in an intermediate position, say at an angle  $\theta$  with the north and south line. Then it may be shown that the current  $C = k \tan \theta$ , where  $k$  is a constant depending on the construction of the instrument and the strength of the earth's magnetic action. Such galvanometers have been much used, but they are liable to disturbance by magnets and masses of iron (gas pipes, steel girders, &c.) in their neighbourhood; and as moreover they have to be adjusted to the N. and S. position, and require some arithmetic to deduce the strength of current from their readings, they are neither so convenient nor so accurate as the commercial ammeters referred to above.

For currents from about a milliampère downwards reflecting galvanometers are always used. The essential point here is that in order to obtain very great sensitiveness, a beam of light is used instead of an ordinary pointer. The usual indicating arrangement is shown in Fig. 72 (plan). The moving part of the galvanometer  $A$ , whether magnet or coil, carries a very light mirror  $B$ ; a lamp  $C$  placed behind a narrow slit in a screen throws a beam of light

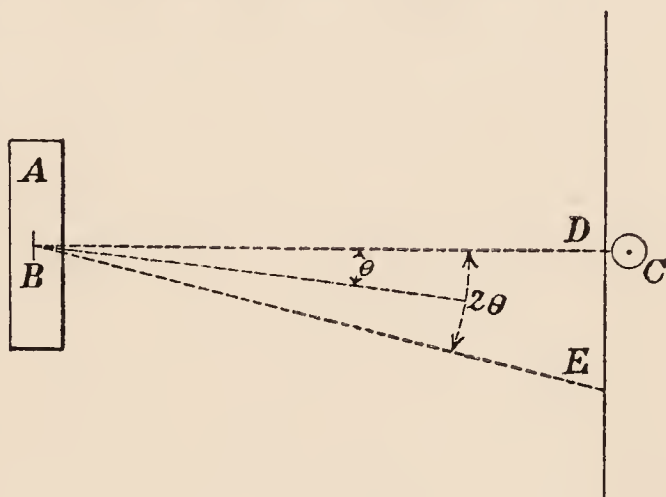


FIG. 72.

on to the mirror, from which it is reflected to form an image of the slit at  $D$  on a scale usually fixed to the screen. The image may be formed either by using a concave mirror, or by causing the light to converge from a convex lens. When electric light is available the filament of an incandescent lamp may itself be focussed on the scale, a more brilliant image being so obtained. When no current is flowing through the instrument, and the mirror is therefore at its normal position, the light is reflected to the middle point of the scale  $D$ ; if now a current produces a small deflection of the moving part of the instrument so that the mirror makes an angle  $\theta$  with its former position, the light will be reflected in a direction  $BE$ , making twice that angle with the incident beam  $BC$ . The scale may be of paper, observed

from the side of the galvanometer, or of ground glass or other translucent material, and read from behind. For the most exact observations it is preferable to replace the lamp and slit by a telescope, and, taking care that the scale is in a good light, observe the point on the scale which appears by reflection in the mirror.

The electrical construction of a mirror galvanometer is essentially the same as that of the instruments already referred to, and consists either of a fixed coil and movable magnet on the plan of the tangent galvanometer (Thomson pattern) or of a coil suspended between the poles of a strong magnet (D'Arsonval pattern). In the former, which until recently at any rate was regarded as the customary type, the magnet is a little bit of steel, perhaps a centimetre long, pasted on the back of the mirror, the whole being suspended by a fine fibre of silk or quartz, so as to offer as little resistance as possible to turning. Two coils are commonly used, placed close together in front and behind the magnet; but in order to increase the sensitiveness still further, an astatic magnet system can be used. This consists of a pair of magnetic needles fixed horizontally at a convenient distance apart on a stiff vertical axis (an aluminium wire); the two needles are parallel, but their poles point in opposite directions, and as they are made of about equal strength, the actions of the earth on the two nearly neutralize, and the combined system has hardly any tendency to set itself one way or the other. The two needles are arranged so as each to lie at the centres of a pair of coils; the astatic system is consequently very easily turned a measurable amount from its position of rest by a current flowing through the coils. As it would be inconvenient to rotate the coils of such a galvanometer in order to bring them parallel to the needle, it is customary to provide a control magnet, or a pair of such, fixed on the stem of the instrument, so that by adjusting them the suspended needle may be turned to suit the position of the coils instead.

By combining the various artifices that contribute to the sensitiveness of the instrument—large number of turns of wire, astatic needle, delicate suspension, mirror method of observation—it has been found possible to make galvanometers that will indicate a current as small as  $10^{-10}$  ampère, or even less. Such a current would have to flow for three centuries to transfer one



coulomb of electricity, and so (according to Faraday's laws) suffice to liberate about a tenth of a c.c. of hydrogen.

For physiological observations a galvanometer of this type is the most suitable, and it may with advantage be wound with very fine wire so as to have as many turns as possible in the coils. This increases the sensitiveness and so allows of measuring smaller currents; it also increases the resistance of the instrument, but, as we have seen, the internal resistance of the muscle or nerve producing the current is so large that the addition of even 10,000 ohms in the galvanometer does not much matter.

The D'Arsonval pattern consists of a coil of fine wire, sometimes circular, sometimes elongated; in either case suspended between the poles of a strong horse-shoe, or c-shaped, magnet. The suspension must be as delicate as possible; it is not so easy to arrange as in the other pattern, since the current has to be led into and out of the coil by the suspending wires; but satisfactory results have been obtained by using a pair of very fine strips of phosphor-bronze close side by side. The coil then offers very little resistance to rotation about the axis of suspension. When a current flows through it, its tendency is to lie with the plane of the coils crossways to the line joining the north and south poles of the magnet; it is therefore necessary to set up the galvanometer with its coils parallel to that line, in order that the least tendency to turn may be observed. A mirror is attached to the coil, and the motion observed with a lamp and scale or telescope and scale in the usual way. A sensitiveness of about  $10^{-8}$  ampères per scale division is usual. The advantages of such galvanometers depend on the fact that the moving part is placed in a strong magnetic field; magnetic changes in the neighbourhood are therefore practically without influence on them, and the instrument may be used even in a dynamo-room, where a Thomson galvanometer would be set in continual fluctuation by the moving masses of iron. As a consequence of this, the sensitiveness, i. e. the current required to produce one scale division deflection, can be determined once for all, and the galvanometer may afterwards be used to measure the actual strength of current in fractions of an ampère. The current is sensibly proportional to the deflection produced on the scale.



*Electromotive force.* There exist, also, instruments for the direct measurement of electromotive force—those of an engineering type being usually called *voltmeters*. When an electromotive force of moderate amount (say  $\frac{1}{10}$  volt and upwards) produced by a battery or dynamo is to be measured with a moderate degree of accuracy a commercial voltmeter with a direct reading scale is appropriate. Such instruments are usually ammeters provided with a resistance coil in series with the working coil; e.g. suppose an instrument be provided of which each scale division corresponds to a current of 0.001 ampère, the resistance of the working coil will probably be a few ohms merely. Let a coil of manganin or other suitable material be put in series with it, making the total resistance 100 ohms; then according to Ohm's law the voltage applied to the whole will be  $0.001 \times 100 = 0.1$  volt per scale division. The instrument with the series coil thus added constitutes a voltmeter, which may be used, e. g., to test the voltage of an accumulator. But such an instrument is only available when the internal resistance of the arrangement producing the electromotive force is negligible. Even a Leclanché cell could not be tested very well with it, for if the resistance of the cell were, say, 3 ohms, the current produced through the voltmeter would not be  $E. M. F. \div 100$ , but  $E. M. F. \div 103$  ohms, i. e. 3 % less; in other words, while the instrument would indicate correctly the voltage applied to it, that will not be the same as the total voltage of the cell, as explained at the end of the last section. If an arrangement with a high internal resistance, such as a muscle-nerve preparation, were tested in this way, the method would break down altogether and the voltmeter indicate practically nothing. If therefore a direct reading instrument is to be used as a voltmeter when the internal resistance is high, it must be one that does not consume any current. These are known by the name of *electrometers*, and are of two types: the 'static' electrometer, usually of what is known as the 'quadrant' (see p. 255) construction, and the capillary electrometer.

The capillary electrometer, due originally to Lippmann, depends on the fact that the surface tension between mercury and an electrolyte such as dilute sulphuric acid, is altered when there is a difference of potential between the two. The usual form of the instrument is shown in Fig. 73. A glass tube of

about 1 mm. internal diameter is drawn out to a very fine bore; this is supported vertically and surmounted by a long glass tube (30-40 cms.), the two being connected by rubber tubing provided with a screw clip. The finely drawn tube dips into a glass cylinder. Mercury is poured on to the bottom of this, and also fills the tube down to a point in the capillary part; the rest of the capillary and the cylinder being occupied by 10 % sulphuric acid. The mercury cannot however flow out of the tube on account of the resistance offered by its surface tension in the very fine bore; the pressure has in fact to be adjusted by altering the height of mercury in the tube (for which purpose the tap at the side is convenient) till the acid-mercury surface is brought to the required point. Electrodes are sealed through the glass to make contact with the mercury in the cylinder and tube respectively; if now an electromotive force be applied by means of these the mercury will move up or down in the capillary. The instrument is mostly used only for null purposes, i. e. to indicate when there is no potential difference between two points, by the absence of movement of the mercury meniscus.

The movements are observed by means of a microscope, and 0.0001 volt can be detected.

When it is desired to measure an electromotive force with accuracy it is best to use an indirect method, that known as the *potentiometer*, in which a galvanometer or electrometer is employed, but only as a null instrument, not to give direct readings by means of its deflection. The connexions are shown in Fig. 74 diagrammatically; *lm* is a thin wire of uniform gauge, commonly a metre long, stretched over a scale; since it is exposed to air, and it is necessary to make contact with it at any point, a hard platinum alloy is the best material to use. Through

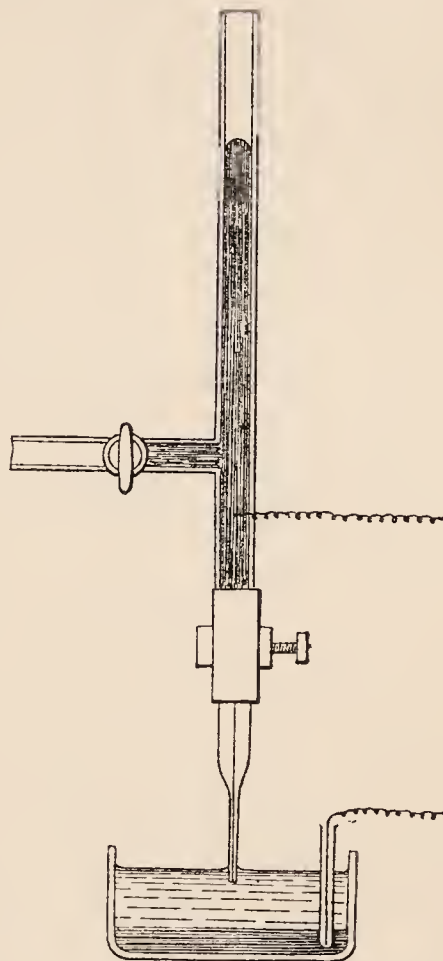


FIG. 73.



this wire a constant current flows from the battery A; this should consist of an accumulator, as it yields more steady currents than any primary cell; the E.M.F. of A must be greater than that to be measured, so that if the latter is higher than 2 volts, two or more accumulators must be put in series. Since a current is flowing through  $lm$  and the resistance of this wire is uniform along its length, the potential falls off uniformly from  $l$  (connected to the +<sup>ve</sup> terminal of A) towards  $m$ , the latter point being about two volts lower than the former. If then the source of electromotive force B to be tested be connected with

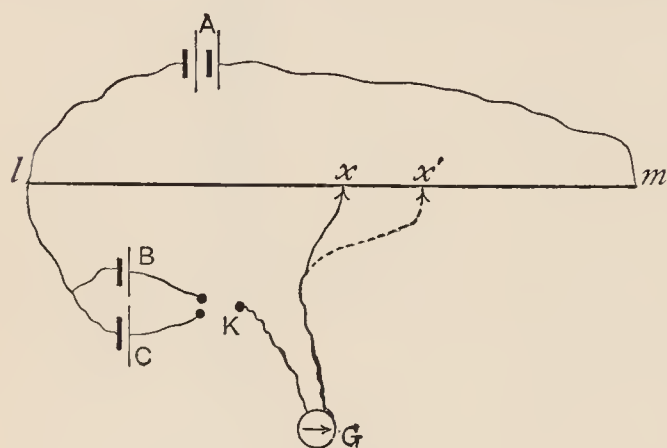


FIG. 74.

its positive pole to  $l$ , there must be some point between  $l$  and  $m$  which is at the same potential as its negative pole; e. g. suppose B to be 1.2 volts, then the point in question,  $x$ , will lie about three-fifths of the way from  $l$  to  $m$ . If therefore the negative pole of B be connected to  $x$  no current

will flow through it; if it be connected to a point on the graduated wire nearer to  $l$  current will flow from B, if to a point nearer to  $m$  current will flow into B from the accumulator. Thus if  $x$  is a sliding contact, and the galvanometer or Lippmann electrometer  $c$  be inserted in the circuit, the point of balance can easily be found. The length  $lx$  then measures the E. M. F. of B. This, however, is on an arbitrary scale; in order to know the actual voltage the experiment must be repeated on a cell of known voltage, such as a Clark c. The key  $k$  allows of putting either B or  $c$  into circuit at will; if the balancing point for the latter be  $x'$ , we have

$$\frac{\text{E. M. F. of B}}{\text{E. M. F. of c}} = \frac{\text{length } lx}{\text{length } lx'}.$$

Measurement of a resistance, like that of an electromotive force, is essentially a comparison with a standard, and is always effected by an indirect method; numerous arrangements of apparatus have been designed for the purpose, but we need only describe two. (i) It follows immediately from the definition



of resistance, that if the E. M. F. applied to a conductor be measured by a voltmeter, and the current by an ammeter, the resistance will be given by dividing the reading of the former by that of the latter. The arrangement is shown in Fig. 75, where *A* is the ammeter, *V* the voltmeter, and *R* the resistance to be tested, *B* the source of current, *K* a key; such a method is suitable for determining the resistance of an incandescent lamp, or any object through which a moderately large current can be led. It must be remarked however that, according to the arrangement shown, any current flowing through the voltmeter must also flow through the ammeter, and so ought to be deducted from the total reading of the ammeter. The voltmeter current is often small enough to neglect, and is zero in a static instrument; but if not very small it should be determined independently in order to apply the above-mentioned correction. Then the volts applied  $\div$  current (corrected) produced = resistance in ohms.

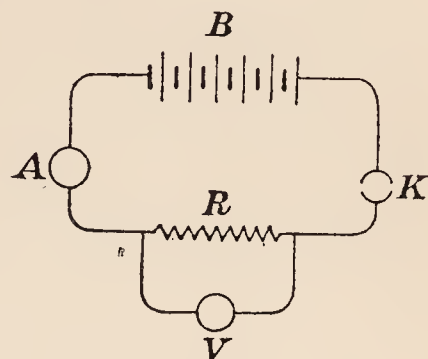


FIG. 75.

When a very high resistance is to be measured—such as a million ohms or more—a modification of this method is available, in which an ordinary mirror galvanometer is substituted for the ammeter, on account of the smallness of the currents involved, and the voltmeter is dispensed with altogether, the current being taken from a battery whose voltage is approximately known.

The most important method for measuring resistances is however the process of comparison known as *Wheatstone's bridge*.

A diagram of this is given in Fig. 76. The current from the battery *B* (usually a single Leclanché cell) divides at *a* into two branches, *PQ* and *RS*, to join again at *c* and return to the negative pole. Two points *b* and *d* in these branches are connected

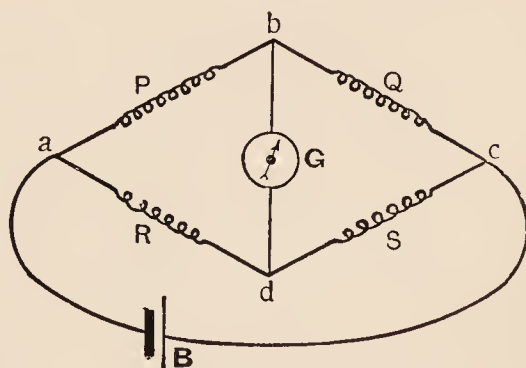


FIG. 76.

together through the galvanometer *G*. If then *P* bears the same ratio to *Q* that *R* does to *S*, there is the same drop in

voltage in passing from  $a$  to  $b$  as from  $a$  to  $d$ , i. e.  $b$  and  $d$  are at the same potential, and no current will flow between them. When this condition is satisfied, then, if three of the resistances be known, the fourth can be found by proportion, say  $s = \frac{QR}{P}$ .

The method is most frequently carried out in practice by the instrument known as a Post-office box (Fig. 77), the parts of which are lettered to correspond with the previous figure. It is not customary to put a terminal at  $a$ , where the current divides, but the battery is connected to  $a'$ , which leads through the intervention of a tapping key to  $a$ , as shown by the dotted line; this

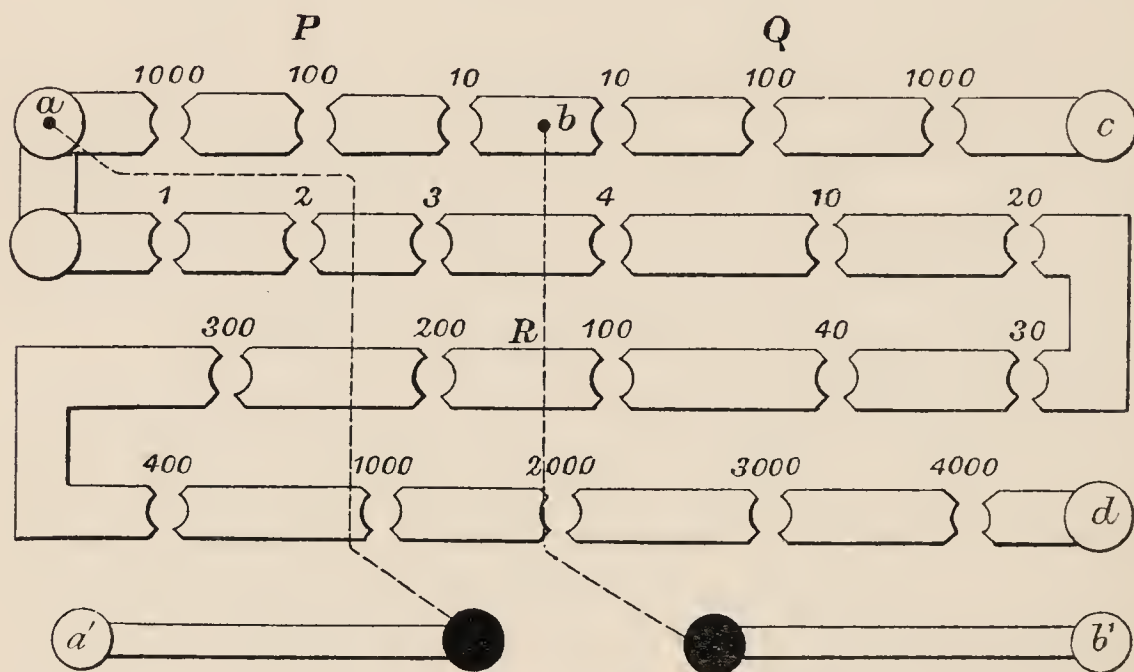


FIG. 77.

key serves therefore for convenient make and break of the battery current. The arms  $P$  and  $Q$  of the bridge are each constituted by a set of three resistance coils of 10, 100, 1,000 ohms, provided with short-circuiting plugs, as explained on p. 208.  $R$  is constituted similarly, but by a set of coils from 1 to 4,000 ohms, such that by taking out appropriate plugs any resistance from 1 to 10,000 ohms can be obtained.  $s$ , the resistance to be measured, is attached to terminals at  $c$  and  $d$ . The galvanometer is connected to  $d$ , but not directly to  $b$ ; like the battery, a special key is provided for it, the terminal of which is at  $b'$ ; by joining this point to the galvanometer that branch of the circuit can be completed,

as shown by the dotted lines, whenever required by pressing the key.

Suppose then it is desired to find the resistance of an incandescent lamp. The lamp is connected to  $c$  and  $d$ ; the 100 ohm plugs are taken out of  $P$  and  $Q$  (the four arms of the bridge should be chosen as nearly equal as possible for greater sensitiveness): 100 ohms is taken from  $R$ , and the two keys put down (the battery key must always be pressed first, to avoid induction effects). It is observed that the galvanometer indicator, the spot of light on the scale, moves away quickly, say, to the left; 1,000 ohms is substituted for the 100 in  $S$ , and then on pressing the keys the spot moves to the right. Then the resistance lies between 100 and 1,000 ohms. By repeated trials we narrow the interval down, and find that it lies between 203 and 204 ohms, say. To get a decimal place the 100 ohms in  $P$  is replaced by 1,000; then, since the denominator of the fraction  $QR \div P$  is magnified ten times, it is necessary to increase the numerator in the same ratio, i. e. to make  $R$  something between 2,030 and 2,040; suppose 2,036 is the amount required to balance, i. e. to produce no movement of the galvanometer, it follows that the resistance of the lamp  $s = \frac{100 \times 2036}{1000} = 203.6$ .

One important class of substances cannot be measured by the ordinary resistance methods, viz. the electrolytes (e. g. blood), for in these the passage of the measuring current would produce decomposition, and so alter the quantity to be measured. It is then necessary to use *alternating currents*, i. e. currents constituted by an oscillation of electricity to and fro along a conductor, instead of a continuous flow in one direction. Such a current would tend to deposit anions and cations alternately on the same pole; these would neutralize one another, and no chemical effect be produced. An alternating current is most conveniently generated in practice by the appliance known as an induction coil, and this is used in place of the battery in a Wheatstone's bridge adapted for measuring electrolytes. The usual physiological pattern of induction coil is not convenient for the purpose however; it should be one of the smallest size, and with a very high frequency of interruption (see description, p. 243), so as to give a clean singing tone when in action; one Leclanché cell should be sufficient to drive it. As a galvanometer gives no indications with alter-



nating currents, it must be replaced by a telephone  $T$ . The apparatus may conveniently be that shown in Fig. 78. The current from the induction coil  $I$  divides at  $a$  between the two parts  $ab$ ,  $ac$  of a wire stretched over a scale; the other arms are  $bd$ , a re-

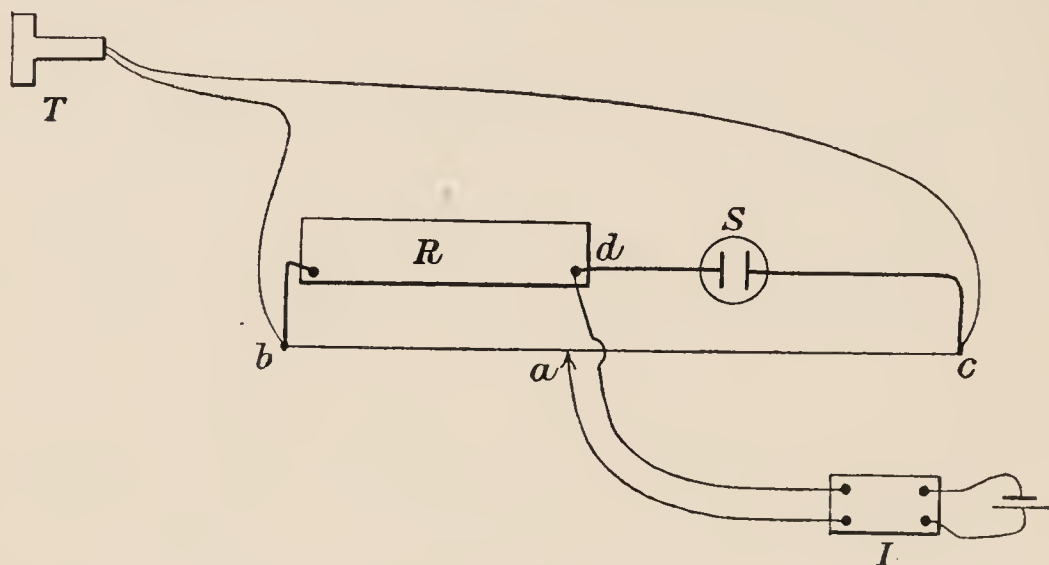


FIG. 78.

sistance box ( $R$ ), and  $dc$  the electrolyte ( $S$ ) to be measured. The balance is obtained by choosing a value of  $R$  approximately equal to  $S$  and then sliding the contact-maker at  $a$  along the wire till the sound in the telephone disappears as completely as possible.

Then

$$S = R \times \frac{\text{length } ac}{\text{length } ab}.$$

The electrolyte is contained in a glass vessel of convenient shape provided with a pair of electrodes of platinized platinum. This is first filled with an electrolyte of known resistivity, such as decinormal KCl, and the resistance taken; then with the liquid to be tested, and the resistance again taken. By proportion the resistivity of the liquid is found. It is customary to express the properties of the electrolyte rather by the conductivity, which (p. 200) is the reciprocal of the resistivity. The conductivity of decinormal KCl at  $18^\circ$  is 0.01119 mho.

## § 6. Electricity and heat.

Certain electrical methods have important application to thermometry and calorimetry, which will be considered here, under two heads: (i) those depending on the production of heat in

a wire due to its resistance ; (ii) applications of the phenomena specially known as 'thermo-electric.'

(i) Since the resistance of a wire increases with rise of temperature, a measurement of its resistance may serve to determine its temperature. The metal most suitable for the purpose is platinum, and the arrangement is commonly known as a *platinum thermometer*. It is made by winding very fine platinum wire on a light frame of mica, enclosed in a tube of glass or porcelain ; the ends of the fine wire are soldered to thick copper or platinum wires, by which connexion is made with an arrangement for measuring the resistance : this is a specially designed modification of Wheatstone's bridge. The resistance of the coil is usually chosen so as to be about 2.5 ohms at  $0^{\circ}$  and exactly 1 ohm greater at  $100^{\circ}$  ; hence every  $\frac{1}{100}$  ohm increase of resistance stands for  $1^{\circ}$  on the scale of the instrument. The scale has been compared with that of a standard gas thermometer, so that the platinum thermometer can be used for exact purposes from the lowest temperatures to more than  $1,000^{\circ}$  c., and is the most convenient instrument which has such wide availability.

Electrical energy is so easily controlled and regulated, that it is often the most convenient means of obtaining a supply of heat. It has been shown that the heat generated in a wire (p. 206) may be expressed as

$$ECt \text{ or } C^2Rt \text{ or } E^2t \div R.$$

From the second of these formulae we see that when a fixed current is led through several conductors (in series) the heat produced in them is proportional to their resistances. But from the third, when several conductors are placed (in parallel) across a pair of mains between which a fixed electromotive force is maintained, the heat generated in each is inversely proportional to the resistance. This is the case with an ordinary electric supply : of the lamps, heaters, &c., wired, those with the lowest resistance get the most current, and consequently the most heat.

If a wire of known resistance be placed in series with an ammeter, and current be passed through it, the heat generated can be calculated from the formula  $C^2Rt$ . The wire may be used to heat the water in a calorimeter, or to serve for any experiment in which a known quantity of heat is required. The heat produced will here be measured in energy units (joules). If at the same time the amount of it be measured by a thermometer,

the relation of the joule to the calorie, i. e. the mechanical equivalent of heat, can be found. Or instead of an ammeter, a voltmeter may be used to measure the electromotive force between the ends of the wire,  $E$  in the formula  $E^2t \div R$ . Or, thirdly, both an ammeter and a voltmeter may be used, and the results calculated by the formula  $ECt$ .

In an experiment on the mechanical equivalent by Schuster the apparatus was arranged as shown in Fig. 78 a. Current from a battery of accumulators  $A$  was taken through the key  $K$ , the heating coil  $H$ , the silver voltmeter  $V$ , and regulated by the rheostat  $D$ . From the ends of the heating coil wires were taken to the battery of Clark cells  $B$  and the galvanometer  $G$ .  $V$  was used instead of an ammeter and a clock, for it is only necessary to know the product  $Ct$  of current  $\times$  time, not these two factors

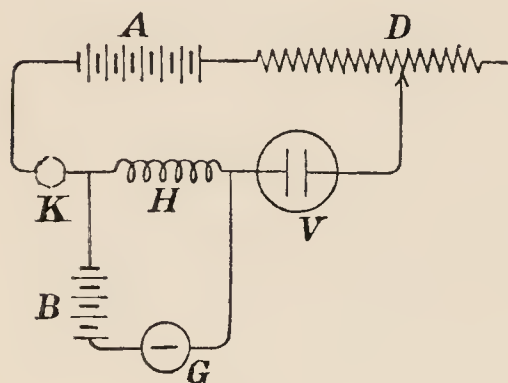


FIG. 78 a.

separately; the product is the quantity of electricity flowing, and is measured by the weight of silver deposited. The Clark cells and galvanometer acted as a voltmeter, in the way described above for the potentiometer (p. 218). In carrying out the experiment  $D$  was regulated so that no current flowed either way through the galvanometer  $G$ . When this is

the case the electromotive force between the ends of the heating coil is precisely equal to that of the battery of Clark cells; hence it is only necessary to multiply the latter (known) amount by the quantity of electricity flowing to get the electrical work spent in joules. The heating coil was placed in a water calorimeter of the usual type, and the heat evolved measured in calories in the usual way. On equating the work spent against the heat evolved, the same result was found as in experiments by the mechanical method (p. 56).

When a circuit is formed out of more than one metal, unless the junctions of pairs of metals be kept all at the same temperature there will be an electromotive force in the circuit. Thus suppose an iron and a German silver wire soldered together at the ends, so as to constitute a closed metallic circuit; let one of the soldered junctions be kept in ice, while the other is heated



to various temperatures ; then it will be found that a current flows. This is the leading observation in *thermo-electricity*. The electromotive force produced is in this case approximately proportional to the temperature ; it is very small—about 0.000029 volt per degree difference in temperature between the two junctions—but can easily be detected by a sensitive galvanometer. The effect is the same if the wires, being directly soldered to form the cold junction, are connected by some other metal—say by a copper wire—provided that the two junctions formed in this way are kept at the same temperature. Thus a thermo-couple is usually constructed after this plan :—

Copper wire		Metal		Metal		Copper wire
from galvanometer	:	A	:	B	:	to galvanometer
		(1)		(2)		(3)

The junctions 1 and 3 being immersed side by side in ice or cold water, junction 2 exposed to a different temperature. The galvanometer, completing the circuit between the two copper wires, gives a reading which varies according to the temperature of the second junction, and hence may be taken as a measure of that temperature. Since however a galvanometer measures current, while it is really the thermo-electromotive force which is characteristic of the temperature, the reading may be erroneous on account of changes in the resistance of the circuit ; it is better, therefore, to take the wires to a potentiometer, and measure directly the electromotive force produced in the circuit. A thermo-couple in conjunction with a potentiometer constitutes an excellent thermometer, available over as wide a range as the platinum resistance thermometer described above. The most suitable pair of metals are platinum and an alloy of platinum with rhodium or iridium.

## CHAPTER VIII.

### ELECTRO-MAGNETISM.

#### § I. Magnetism.

A MAGNET shows in the most marked way the phenomenon of polarity, i. e. it possesses, essentially, ends which show precisely opposite characters. It is well known that if a magnet be suspended or pivoted (like a compass needle) one end will point towards the north, the other to the south ; and it is found that if a magnet be cut in two, the two points will not be, one north-seeking, the other south-seeking, but each will be a small complete magnet, possessing polarity like the larger one. And this is true, however small the piece of magnetized matter may be, even if it be merely iron dust, so that one must look upon the polarity as an essential character, in fact suppose that each molecule of a magnet is a magnet itself, and the magnetic effect of the whole is due to the molecular magnets being arranged in such a way as to assist one another.

The effects of magnets may be most conveniently explained by means of what is known as the *field of force* round them. Anywhere near one magnet another one tends to set itself in a definite direction ; thus if a bar magnet be placed on a table and iron filings be sifted over the table, they will lie with their long axes in definite directions (to prevent the filings being drawn bodily to the big magnet a sheet of glass may be put between). Each filing is made into a small magnet, and points with its north pole in the direction of the field due to the large magnet ; hence the appearance of the filings will give a picture of the field, showing its direction at different points. Some diagrams obtained in this way are given below, and should be carefully studied ; thus Fig. 79 shows the arrangement round a simple bar magnet ; the filings are seen to lie in continuous lines stretching from the positive or north pole

round through the air to the negative, or south pole ; the curves so marked out are called *lines of magnetic force*, because the magnetic force acts everywhere along their length. Such lines

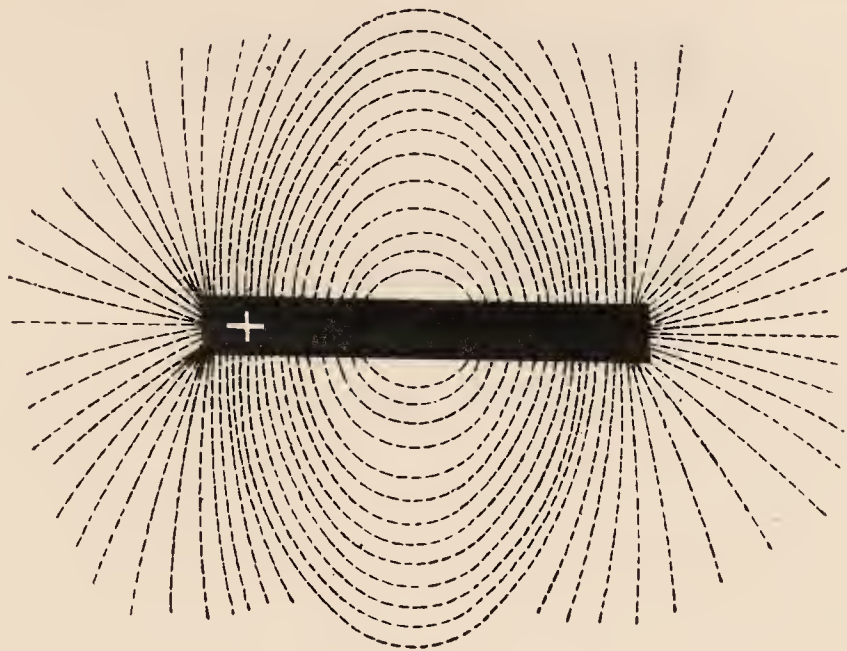


FIG. 79.

always proceed from a N. to a S. pole—not necessarily that on the same magnet, as may be seen from Fig. 80, which is obtained with two bar magnets placed with their unlike poles



FIG. 80.

opposite each other. In Fig. 81 the N. poles of the two magnets have been put close together, and it is seen that the lines of force produced repel one another, all being driven round to the distant S. poles.

From this behaviour of the lines of force, there results an attraction between unlike magnet poles (N. and S.), a repulsion between like poles (two N. or two S.); it is in fact as if there were a tension along the lines of force, a pressure at right angles



to them; if Figs. 80, 81 be thought of with this pressure and tension in mind, it will be seen that there is a tendency for the poles to approach in the former, to separate in the latter case.

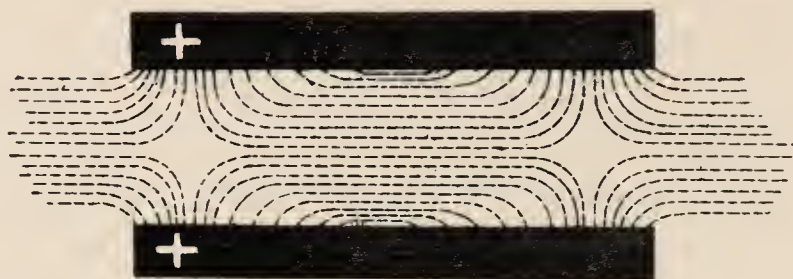


FIG. 81.

By means of these forces we get a definition of the strength of a magnet pole (or *quantity of magnetism* located there). It is found that the mechanical force between two poles is proportional to the product of their strengths, and inversely proportional to the distance between them. Hence we may define the unit pole as that which, placed 1 cm. from a similar and equal pole, will repel it with a force of 1 dyne (ignoring the influence of the other ends of the magnets); the law of force may be stated algebraically as

$$f = mm' \div r^2$$

where  $m$ ,  $m'$  are the strengths of the poles,  $r$  the distance apart, and  $f$  the force exerted.  $f$  is positive (repulsion) if  $m$  and  $m'$  are both positive (N.) or negative (S.), it is negative (i. e. an attraction) if  $m$  and  $m'$  are of opposite signs.

The fact that a magnet cut in two forms two magnets, may now be repeated in more quantitative form; a bar magnet possesses a pair of poles, at its ends, and the two are necessarily of equal strengths, one consisting of  $+m$  units of magnetism, say, the other of  $-m$ . If the bar be cut in two the two cut ends will themselves become poles of the same strength, so that each half of the bar will now have  $+m$  units at one end,  $-m$  at the other; if the two halves were fitted together in their original position, these two new poles would neutralize, for the effect of  $+m$ ,  $-m$  units at the same spot is obviously zero. The total magnetic effect of the magnet is best measured by multiplying the strength of either pole into the distance between the two; the product so obtained is called the *magnetic moment*. Then if a bar magnet be cut into two equal parts, although there is the same quantity

of magnetism on the ends as before, each small magnet will have only half the magnetic moment of the original magnet, which is an appropriate expression of the fact that there is only half as much magnetized matter in it. Accordingly the magnetic moment per cubic centimetre of metal serves to measure the effectiveness of the material in producing magnetic action, and is called the *intensity of magnetization*.

A good quality of soft iron can attain some 1,200 units of intensity of magnetization, steel about 800-1,000; thus a bar magnet of 1 sq. cm. cross section, and 10 cm. long, might have a magnetic moment of 8000, and consequently possess +800 units of magnetism on one end, -800 on the other.

The strength of a magnetic field (also called the *magnetic force*) at a point acts, as we have seen, in a definite direction, so that it belongs (like force, acceleration, &c.) to the class of quantities known as vectors (p. 26); in order to express it we must state both its magnitude and direction. The magnetic force at any point is defined by the *mechanical force which would be exerted on a unit positive quantity of magnetism placed there*. Hence, if a quantity  $m$  of magnetism be put at a place where the magnetic field has a strength  $H$ , the force exerted on it is  $mH$ .

When a magnet is placed in a *uniform* field, there is no tendency to move it bodily, but only to rotate it into the direction of the lines of force. Thus (Fig. 82) if the magnet  $AB$  lie at an angle  $d$  with the lines of force, there is a mechanical force on the north pole  $A$  represented by the arrow, and in accordance with the definition of the strength of magnetic field it will amount to  $mH$ ; there is an equal and opposite force on the south pole at  $B$ ; these two constitute a couple whose moment is = either force  $\times$  perpendicular distance between their lines of action, i. e.  $mH \times l \sin d$ . If we call  $ml$  the magnetic moment =  $M$ , the couple is  $MH \sin d$ . When  $d = 0$  this vanishes, i. e. when the magnet lies along the direction of the magnetic field there is no longer a couple on it, and it is in equilibrium.

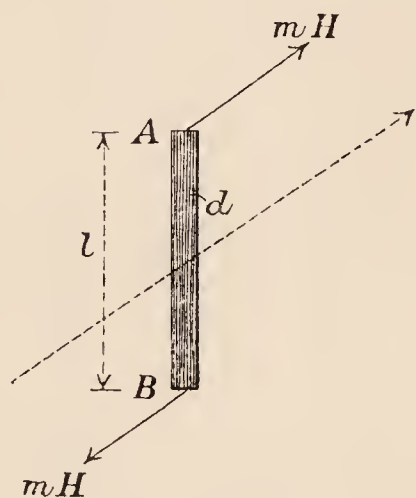


FIG. 82.



When a magnet is placed very near to another magnet, it suffers a force tending to move it bodily towards the latter. This is due to the field near the magnet being far from uniform, and at very moderate distances the bodily attraction becomes inappreciable.

As an instance of a magnetic field that of the earth may be studied; the earth behaves as a magnet, and consequently at any point near it there is a field of magnetic force, disposed in somewhat the same manner as the field round the bar magnet of Fig. 79. Only on account of the enormous size of the earth, the strength and direction of the field within any room is practically constant; hence we have the problem of determining the field at any desired spot (a magnetic observatory). Now to specify a vector three quantities must be stated—there are, as has been remarked, various ways of doing it—and the three that are commonly chosen are as follows:—

(1) A vertical plane can be imagined containing the actual direction of the field; the angle this plane makes with the geographical north is called the *declination*. In London at the present time (1902) it is about  $17^\circ$  to the west. By this angle the vertical plane is defined. The direction itself is called the magnetic meridian.

(2) The actual direction of the magnetic force within that plane makes an angle with the horizontal which is called the *dip*. This is at present about  $67^\circ$  downwards towards the north. By these two angles the direction of the field is completely defined.

(3) Within the given vertical plane the magnetic force (like a mechanical force) may be resolved into a horizontal and a vertical component; the magnitude of either of these serves, together with the dip, to determine the other, and consequently the total force. It is the horizontal component that is usually stated; it is about 0.183 units, now, in London.

The direction and intensity of the earth's magnetic field varies in passing to different parts of the earth's surface; this is as we should expect, for in doing so we become differently situated with respect to the magnetic masses within the earth. There are two points near, but not at the geographical poles, where the angle of dip is  $90^\circ$ , i.e. a magnet points vertically downwards; these points must evidently lie over the magnetic poles. About



halfway between them lies a line, roughly a great circle of the earth, at any point of which the angle of dip is  $0^\circ$ , i.e. the lines of force are in the horizontal plane; this is the magnetic equator.

In a similar way, though on a much smaller scale, the field round an ordinary magnet might be traced out.

To determine practically the magnetic field at a point, the following experiments are required :—

(1) A suspended or pivoted magnet (such as a compass needle) free to move in a horizontal plane will clearly indicate the direction of the horizontal component, and so if the true north be known, the declination can be measured.

(2) A magnet pivoted so as to be free to move in a vertical plane (called a dip-needle) is placed in the magnetic meridian : it will then lie in the direction of the actual magnetic force, and the angle of dip may be measured.

(3) A magnet is placed with its axis east and west, and a compass needle is put in line with it, a short distance away (Fig. 83, the magnet is said to be end on to the needle). The needle is then deflected out of its north and south direction through a certain angle, say  $d$ . If  $M$  is the moment of the magnet (AB in the figure), and  $r$  the distance between its centre and that of the needle,

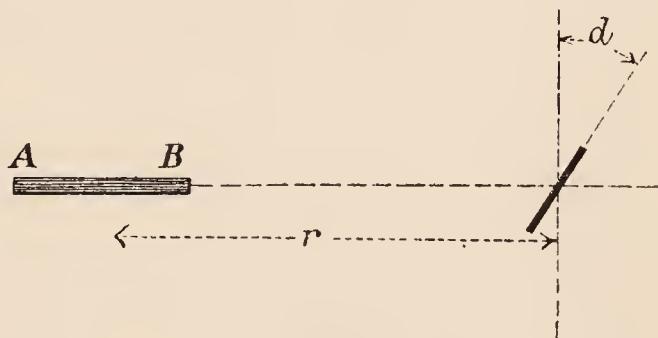


FIG. 83.

$$H = \frac{2M}{r^3 \tan d}$$

(4) The same magnet (AB) is suspended and allowed to oscillate. If  $T$  be its time of vibration,  $l$  its length,  $b$  its breadth (the magnet being supposed rectangular),  $w$  its mass,

$$H = \frac{\pi^2 (l^2 + b^2) w}{3 MT^2}$$

By combining these two observations the necessity for an independent measurement of  $M$  is avoided.

When a piece of iron or steel is placed in a magnetic field it becomes magnetized, or magnetism is induced in it. It would be more correct to say that the magnetism exists beforehand in the iron, the individual molecules being magnets, but that

ordinarily the magnetized molecules are scattered at random so that they produce no more magnetic moment, on the whole, in one direction than in another. What the magnetic field accomplishes is to turn the molecules prevailing in one direction, so that they add their effects together, and produce an appreciable effect on outside points.

A magnetic field is most conveniently produced by means of an electric current, as explained in the next chapter. Without going into details here, we may assume that we have a means of exposing a long thin bar of iron to a field, and increasing the intensity of the latter step by step as desired. If the bar be of soft (wrought) iron, these phenomena are observed:—At first the intensity of magnetization produced is proportional to the strength of field, and may be perhaps ten times as great; this is only true, however, while the field is very small, not more than about 0.04 unit. After this the intensity of magnetization increases more and more rapidly till, when a field of four or five units is applied, the magnetization may be 100 or more times the field; this is due to group-

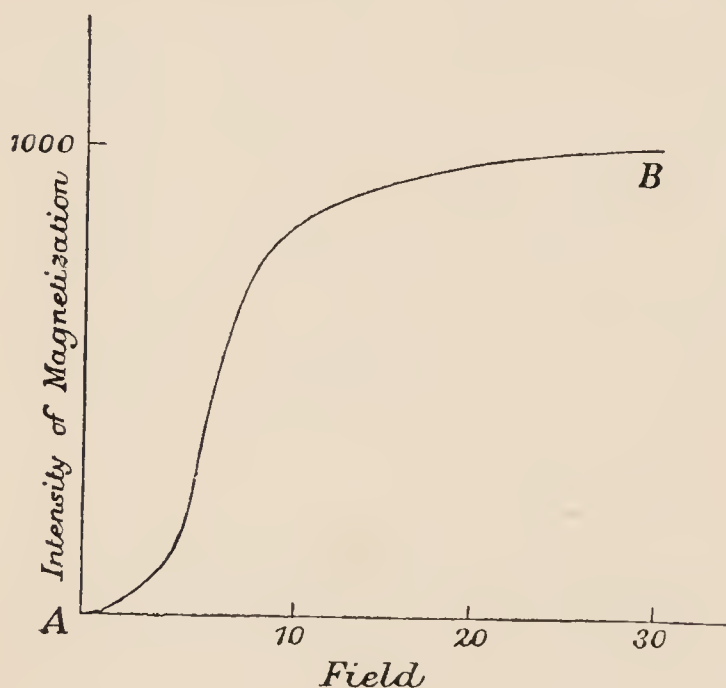


FIG. 84.

ings of molecules being broken up, and the molecules suddenly swinging round till nearly parallel to the magnetizing force. If the field be still further increased, this process of orientation proceeds further, but it must obviously reach a limit when all such molecular groups have been acted upon. Accordingly, with a field of 20 units an intensity of magnetization of about 1,000 is reached, the ratio  $\frac{\text{magnetization}}{\text{field}}$  having

fallen to  $1000 \div 20 = 50$ .

Beyond this, an increase in magnetizing force produces but little effect—a field of the enormous strength of 2,000 units has only induced in the best specimens of iron an intensity of magnetization of about 1,700; the iron is said to be saturated. The course of these observations is reproduced in Fig. 84, line AB.

If now the magnetizing force be reduced, the magnetism will not disappear: even when the force is removed entirely, some 80% to 90% of

the magnetic moment may remain, the metal being said to possess *retentivity*. This indicates that the magnetic molecules, once forced into position with their axes parallel, do not spontaneously fall back into an irregular arrangement when the force ceases to act. Consequently we have brought the specimen back, so that it is exposed to the same conditions as at first (no magnetic force), and yet it is not in the same internal state as at first ; the magnetization of iron consequently depends not only on the conditions it is subjected to at the time, but on its previous history as well. This phenomenon, which is known as *hysteresis*, is shown by solids in other ways (p. 138).

Both soft iron and steel show retentivity, the former to even the greater extent of the two. To reduce the magnetization to nothing again, it is necessary to apply a magnetizing force in the opposite direction ; or else to use other means, such as shaking, hammering, raising to a red heat, &c. The main difference between soft iron and steel is that it is very much easier to remove the residual magnetism from the former. Consequently steel is used for making *permanent magnets*, i. e. such as will stand ordinary usage without losing their magnetism ; while soft iron is used for making *electromagnets*, which are intended to be magnetic whilst a current is flowing, and to lose their magnetism when it stops. For details as to the construction of the latter, see p. 236, below.

## § 2. Magnetic field due to a current.

An electric current produces a magnetic field in its neighbourhood. This fact may be demonstrated by the aid of iron filings, just as in the neighbourhood of a steel magnet. If a long vertical wire be arranged with a sheet of cardboard placed horizontally round it, and filings be sifted on to the card, then, when a current is passed through the wire, the filings will set themselves along the lines of force, and these will be found to be circles round the wire. (The rest of the circuit through which the current passes should be kept well away from the card.) The same fact may be shown by a compass needle : if placed near such a wire (carrying a strong current) it will set tangentially to a circle drawn round the wire : if taken further away, the influence of the current, though always the same in character, will be less, so that the directive force of the earth will make itself more and more felt. The direction taken by the north pole of the magnet, in such an experiment, indicates the sense in which the current is flowing according to Ampère's rule (p. 188).



If, as usually occurs in practice, the out and return wires of a circuit run near together, there is a strong magnetic field between the two wires, but a weak one outside. The reason for this will be seen from Fig. 85, in which the wires are supposed

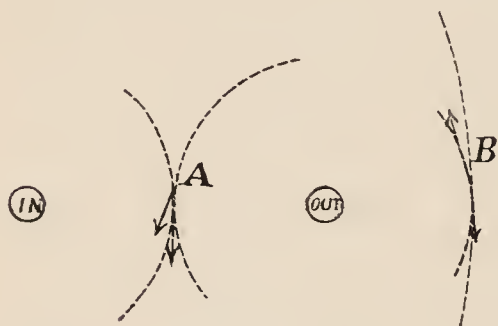


FIG. 85.

to lie perpendicularly to the plane of the paper: the arrows indicate the field produced by each current separately, and it will be observed that at A the direction of these is such as to help one another, but at B to oppose. Hence, by twisting together the wires carrying equal currents in opposite directions, their magnetic action on external points

may be practically avoided. This is a rule that should be attended to when currents of any strength are employed in the neighbourhood of a delicate Thomson galvanometer.

In the construction of a galvanometer, the object here being to make the magnetic action of the current as large as possible, the wire is wound in a circular coil, often of thousands of turns, and the 'needle,' or indicating magnet, is placed at the centre. In this way the effects of all the small portions of the wire in the coil are in the same direction, and help one another. For, fixing attention on some point in the coil, a plane at right angles to the wire at this point will pass through the centre of the coil: it will in fact be a radial plane. A circle drawn round the point, in this plane, and so as to pass through the centre of the coil, will, at the centre, be tangential to the axis of the coil, and consequently the current near the chosen point will cause a magnetic field along the axis: but so will all the other parts of the wire, for the same reason. (This may be clearly seen with the aid of a simple model.)

Now it is found that the magnetic field produced by a short element of current at a point in the plane at right angles to the current, is proportional to the strength of current: and to the length of the element considered: and inversely proportional to the square of the distance of the point from the current.

Accordingly the following definition of unit current is convenient, and is adopted, viz. that *current which, flowing along*

1 cm. of wire bent into an arc of a circle, so as to be all at a distance of 1 cm. from the centre, produces unit magnetic field there. This is known as the centimetre-gramme-second (C. G. S.) unit of current; the ampère, however, is only  $\frac{1}{10}$  as much.

In a galvanometer, if there be  $n$  turns of radius  $r$  the length of wire will be  $2\pi nr$  cms., and as each point is at a distance  $r$  from the centre, the magnetic field produced by a current of  $\iota$  (C. G. S. units) is

$$\iota \times 2\pi nr \div r^2 = \frac{2\pi n\iota}{r}$$

If  $\iota$  (as usual) be in ampères, the field is

$$\frac{\pi n\iota}{5r}$$

If then a tangent galvanometer be arranged, as described on p. 212, with its coil in the magnetic meridian, the field at the centre is (1)  $H$  towards the north, where  $H$  is the strength of field due to the earth, (2)  $\pi n\iota \div 5r$  towards the east or west according to the way the current flows. The resultant of the two is along the diagonal of a parallelogram constructed with these two magnetic forces as sides, and consequently makes with the north and south direction an angle  $\theta$ , such that

$$\tan \theta = \frac{\pi n\iota \div 5r}{H}$$

The magnet needle lies along the direction of the resultant magnetic force, consequently  $\theta$  is the angle through which it is turned, and  $\iota = 5rH \tan \theta \div \pi n$ . The current is thus proportional to the tangent of the angle of deflection  $\theta$ . If  $H$  be 0.18 unit this becomes  $\iota = \frac{0.29r \tan \theta}{n}$  ampères.

In order to magnetize a bar of iron, it is convenient to use a coil wound in close spiral form, of a length about equal to the bar. By this means a field is obtained which is directed along the axis of the spiral, and is uniform along the length of it, except near the ends. It may be shown that the strength of field inside a coil of  $n$  turns forming a spiral of length  $l$  is  $= 4\pi n\iota \div l$ , where  $\iota$  is the current in C. G. S. units. If the

current be expressed in ampères, we have consequently field  $= 0.4 \times \pi n i \div l$ . In this case the field inside the coil is independent of the radius of the cylinder, provided only it be long compared with its breadth. Outside such a coil there is hardly any magnetic field.

For example, in order to produce a field of 10 units, which is nearly enough to saturate ordinary soft iron, we might wind a coil of wire 1 mm. in diameter (including the insulation) and have 10 turns to the centimetre. If only one layer of windings be used and 0.796 ampère run through the coil, the field inside is  $0.4 \times 3.1416 \times 10 \times 0.796 = 10$ , the required amount. Such a current might be allowed to flow for any length of time, without making the coil unduly hot. A very much stronger current might be used for a short time, and so a much stronger field produced temporarily, and as the magnetism is induced almost instantaneously, it would suffice for the object in question.

In order to find which is the north pole of a bar magnetized by a current, we may adapt Ampère's rule, and say 'swimming in the current and looking towards the magnet, the north pole will be that end of the bar which lies towards the left.' Put in another form this amounts to 'the direction of the current and the direction of the magnetic field (S. to N. through the iron) are related as the rotation and translation in an ordinary screw.'

An *electromagnet* consists of a core of soft iron, usually either straight or bent into the shape of a horse-shoe, wound over with a magnetizing coil. A short distance from one or both poles is placed a piece of iron called the 'armature,' usually supported by a spring, so that it can be moved to and from the poles. When a current is passed the iron becomes a magnet, and attracts the armature. When the current is turned off, although it is possible for soft iron to retain a large fraction of its magnetism, when carefully handled, practically the shock destroys it all, so that the attraction ceases and the spring draws the armature back again. In the electromagnetic tuning-fork described on p. 4 the prongs of the fork act as armatures. Another example of the use of an electromagnet occurs in the ordinary pattern of induction coil.



## § 3. Mechanical force on a conductor.

When a conductor carrying a current is placed in a magnetic field, it suffers, in general, a mechanical force. The amount of this may best be expressed by considering a short length of the conductor by itself. Let this length be  $l$  (Fig. 86), and the current flowing through it  $i$ , and let the field  $H$ , to which it is exposed, be of the magnitude and direction indicated by the arrow in the figure. Then the force acting on the conductor is measured by the area of the parallelogram, whose sides are  $li$  and  $H$ , and it is perpendicular to the plane of the parallelogram. Hence when the conductor lies along the direction of the magnetic field, there is no mechanical force on it, for the parallelogram vanishes in this case; when the current is at right angles to the field the force is greatest. The direction of the force is at right angles to both the current and the field, but in order to state in which sense it acts, a separate rule is needed, which is in fact the converse of Ampère's rule. It may perhaps be put most simply in this form; if a current flow vertically *up* a wire, and be placed in a magnetic field directed towards the *north*, it will be forced towards the *west*; if either the field or the current be reversed, the mechanical force is reversed: consequently, if they are both reversed, the force remains in the same sense.

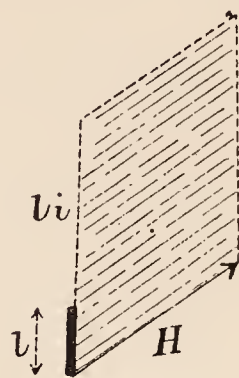


FIG. 86.

The mechanical action on a complete circuit may be obtained by compounding the forces on all the short lengths, or elements, of which it is made up.

It is on the forces just described that the action of electromotors depends, and we may take instances of the construction of electromotors to illustrate the laws stated above. The simplest form of motor (one of no practical use) is Barlow's wheel: this consists of a metal disc, capable of rotation on a spindle, provided with two contacts for leading current in and out, one on the spindle, the other on the edge of the disc; the contacts may be made by flexible 'brushes' of copper, as in dynamos, or, to reduce friction, by mercury cups. A strong magnet is placed so that it produces a field transversely to the plane of the disc. Let us suppose the spindle placed horizontally and pointing N. and S., and the magnet arranged so that the field across the disc is directed towards the north; suppose the current led in by a brush touching the lowermost point of the disc, and out by the spindle; then it flows vertically upwards across the disc, and, according to the preceding rule, there will be a force acting on the lower part of the disc (through which alone the current flows) urging it towards the west. This force will cause the disc to rotate, and as the rotation brings constantly fresh points of the disc into contact with the

brush, and so keeps the line of current in the same absolute position in space, the rotation will be kept up indefinitely.

In an ordinary electromotor the rotating part, or armature, may consist of a drum wound over with wire, each turn of wire proceeding along one side of the cylinder, across the end face, and back by the other side; the ends of each turn, or each small group of turns, are brought to a 'commutator,' which consists of copper bars arranged lengthways on the spindle of the drum. The object of this is to allow the current to be led by brushes through half the armature segments in one sense, and the other half in the opposite sense; and, as the rotation takes place, the current is transferred from one section to another, but the geometrical position of the sections carrying the current in either sense remains, as in Barlow's wheel, the same. A very strong magnetic field (usually produced by an electromagnet) is arranged crossways to the spindle of the armature. For simplicity of description we will suppose the field, as before, to be directed towards the north; the spindle to be

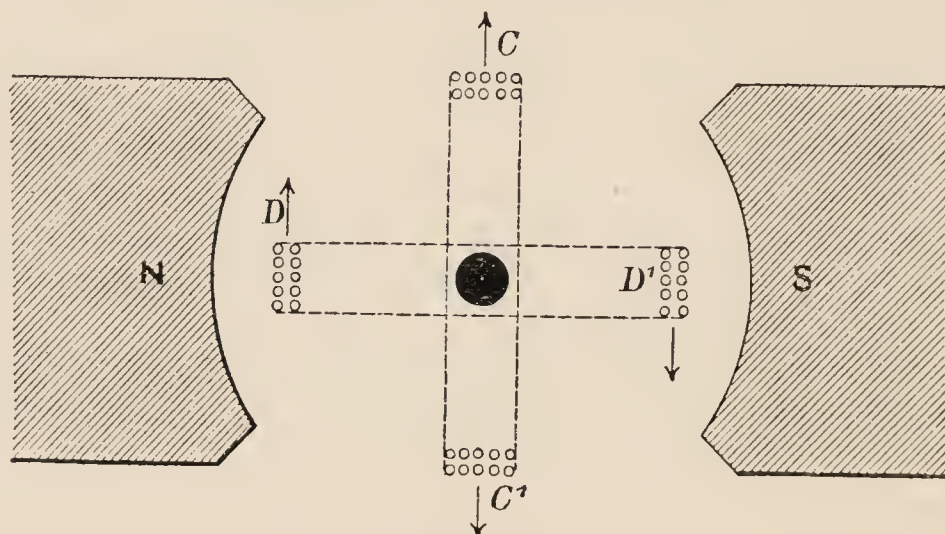


FIG. 86 a.

horizontal, and consequently to lie east and west. Consider now a section  $cc'$  (Fig. 86 a) of the armature which lies in a vertical plane: if a current be led through this, there is a force on the conductors lying east and west, but since it must be at right angles to the conductor, and also to the field, it must be a vertical force, and therefore act through the centre of the spindle and produce no effect in rotating the armature; such a segment of the armature is, therefore, without action. Next consider a segment  $dd'$  lying in a horizontal plane; the forces are again vertical, for the same reason as before. They of course do not act through the spindle; the part of the conductor which is carrying current from east to west suffers a force vertically downwards, the opposite part through which the current flows from west to east suffers an upward force, and the resultant of these two is a couple tending to rotate the armature. Hence the brushes must be arranged so as always to lead the current through that section of the



armature which lies horizontally. These relations should be studied with the aid of a model. The diagram (Fig. 86a), which is supposed to be drawn at right angles to the spindle, illustrates them partly.

Another good illustration of the principles of this section is to be found in the D'Arsonval type of galvanometer. This, as already explained, consists of a coil of wire through which the current passes, suspended between the poles of a strong magnet. The coil is arranged so that its plane coincides in direction with the lines of magnetic force. The latter are horizontal, and consequently any portion of the wire of the coil which lies vertically experiences a mechanical force at right angles to itself and to the field. If, to distinguish the various directions involved, we suppose ourselves looking at the face of the instrument, that the magnetic field across the coil is from left to right, and that the current is flowing upwards in the left-hand side of the coil, downwards in the right, it follows from the rule that there is a mechanical force driving the left-hand side of the coil inwards, the right-hand side outwards; these constitute a couple, and will rotate the coil; but not indefinitely, as in the electromotor, on account of its movable contacts. In the galvanometer, the rotation will only take place so far that the balancing couple due to the twist of the suspending wires is equal to that applied electromagnetically; and since, when the angle through which the coil turns is small, the couple exerted by the suspension is proportional to it, the amount of turning is proportional to the electromagnetic couple, and hence to the current flowing through the coil. If the coil were quite free to turn, it would not rotate continuously, but only through  $90^\circ$ , for then it would come into a position similar to that of the inactive motor segment described above; the forces on the two sides of the coil would be in opposite senses, but along the same straight line through the centre of the coil, and would neutralize; the coil would then be in a position of rest. If the current be made to flow through the coil in the opposite sense, the magnetic field (produced by an independent magnet) remaining the same as before, the sense of the rotation will also, obviously, be reversed.

#### § 4. Induction of currents.

We saw, in the last section, that when a conductor carrying current is placed in a magnetic field, there is a mechanical force tending to cause it to move. The converse phenomenon is also well known: when a conductor is moved in a magnetic field, an electromotive force is produced in it, so that if it forms part of a complete circuit, a current flows through it.

The mechanical force, as we have seen, is such as to cause the conductor to move at right angles to the direction of the magnetic field; and



accordingly, in the converse case of induction, it is only so far as the motion is in that direction that it is effective. For definiteness we will suppose that the conductor is vertical, and the magnetic field towards the north. Then, in order to induce currents in it, the conductor must be moved east or west (i. e. must have a component movement east or west—it may go slantways, but the components in other directions are ineffective). If it be of length  $l$ , and the field be of strength  $H$ , the electromotive force due to a motion of 1 cm. will be proportional to  $lH$ : we may here conveniently make use of the term magnetic induction as expressing the magnetic flux—somewhat vaguely expressed by ‘lines of force’ hitherto, and say that  $lH$  is the amount of magnetic induction cut by the conductor in moving one centimetre to the east or west; and if it have a velocity  $v$  cms. per sec., it will cut  $vlH$  of magnetic induction per second. With the aid of this conception it is easy to state the law of the induction of currents as follows: ‘The E. M. F. induced in any conductor is equal to the quantity of magnetic induction cut by it per second.’ Again, as to the sense in which the induced currents will flow, we have what is known as Lenz’s law, which is merely a simple deduction from the conservation of energy: the induced current is always in such a direction that its reaction tends to stop the motion inducing it. If the contrary were true, a motion of a conductor in a magnetic field, once started, would cause a force tending to increase itself, and so would accelerate indefinitely, and an indefinite amount of work might be done without any expenditure of energy, which is absurd. Hence, reverting to the case imagined above, if the wire be moved from west to east, an upward current will be induced in it, since, as we saw in § 3, an upward current in the wire would cause a force towards the west, and so oppose the motion.

One of the simplest pieces of apparatus that illustrates these phenomena is the *earth coil*. This consists of a good-sized flat coil of wire wound on a round or square frame (without iron in it), arranged so that it can be laid horizontally, and quickly turned over face for face. When it is being turned over an electromotive force is induced in it, and, if the ends be attached to a galvanometer, a current will be indicated. The induction here is on account of the vertical magnetic field of the earth. If  $V$  be the strength of that field,  $A$  the area of the coil, and  $n$  the number of windings, the total magnetic induction through it is  $nAV$ ; when it is turned over this must be regarded as becoming  $-nAV$ , since it enters the opposite face of the coil. There has, therefore, been a change of  $2nAV$  in the magnetic induction. The current at any instant is found by dividing the *rate of change* of the induction by the resistance of the circuit: this would vary according to the speed with which the inversion was carried out; but it is not necessary to measure it, for the total quantity of electricity that flows through the circuit is found by dividing the total change of induction by the resistance or  $Q = 2nAV \div R$ . If the motion be effected

quickly this will give a single throw to the galvanometer needle, the amount of which is proportional to  $Q$  (a galvanometer so used is called *ballistic*). Hence an earth coil can conveniently be used to calibrate a ballistic galvanometer.

These phenomena can be very well illustrated by the same instances as were used in the last section, the D'Arsonval galvanometer and the electromotor; for an electromotor, run by some external power, constitutes a dynamo, and generates current in the reverse direction to that needed to drive it as a motor.

Consider first, in the motor, a segment of the armature in the position which we saw is ineffective in producing rotation; if it be rotated by external power, then at the moment of passing through this position it is ineffective in producing current; for both the upper and lower conductors are then travelling in the direction of the lines of magnetic force, and consequently do not cut them: the magnetic induction passing through the circuit is for the moment constant, and there is no induced electromotive force.

On the other hand, a segment which lies in the direction joining the poles of the magnets is moving up at one side, down at the other, and cutting the flow of magnetic induction rapidly: there is consequently an induced electromotive force in it; and the brushes are so arranged as to make contact at any moment only with those segments whose rotation is at that moment effective; so that a constant electromotive force exists between the brushes, and may be used to supply an external circuit.

Such a device is called a dynamo; the magnetic induction is produced by electromagnets, kept in action by a part of the current from the dynamo itself. Although this involves the consumption of a certain part of the power of the machine, in merely keeping up the magnetic field, it is commercially more advantageous to do so than to use permanent steel magnets, because the latter are not so strong, and consequently it is not possible to get so much power out of a machine of a given size. But for medical coils, this objection is of no weight, and as a machine with permanent magnets has fewer parts to go wrong it is commonly preferred, and is distinguished by the name of 'magneto.'

We have seen that the electromotive force induced in a conductor depends only on the rate at which it cuts the magnetic induction; in no way upon the thickness of the conductor itself. If, therefore, the strength of field and the speed of rotation be maintained constant the electromotive force generated will be proportional to the number of turns in the armature coils which are in circuit at any moment. Hence a dynamo that is intended to give a high voltage is wound with comparatively thin wire; one for low voltage, but large current, with thick wire, so as to lessen the internal resistance. It must be remembered that the product volts  $\times$  ampères measures the electrical power given out, and that this



depends essentially on the power of the engine used for driving the dynamo.

The D'Arsonval galvanometer shows with equal clearness the reciprocal relation between the force on conductors in a magnetic field and the induction of currents. When current is passed through such a galvanometer, the coil tends to set itself in the plane at right angles to the lines of magnetic force. Suppose a deflection thus given to the coil, and thereafter the circuit to be broken; the coil will swing back to its equilibrium position, and oscillate to and fro. In doing so, however, it cuts the lines of force—the amount of magnetic induction through it is constantly changing, and, therefore, there is an induced electromotive force in the coil; if the coil is on open circuit no current will flow, and no particular effect will be observed: the oscillations will probably last for a considerable time, indeed the galvanometer may be a satisfactorily 'ballistic' one. But a galvanometer is often used with a key that short-circuits it when not in action: suppose that to be the case. Then the induced electromotive force will generate a current in the coil, and, according to Lenz's law, this will be in such a direction as to stop the motion; the induced current, therefore, damps the oscillations, and may even do this so effectually as to make the instrument dead-beat, that is return to its position of rest without oscillating.

We have so far considered induction of currents only as a consequence of motion of a conducting circuit in a magnetic field. It is a matter of indifference whether the field be due to permanent magnets or currents or both: further, it is a matter of indifference whether the circuit or the magnets move: if the armature of a dynamo were kept fixed and the field magnets caused to rotate round it, the action would be precisely the same; similarly in an ordinary galvanometer the vibrations of the suspended magnet needle may be damped by the reaction of induced currents in the coil, only on account of the very weak field that it has the effect is small.

But all that is really of consequence for the production of induced currents in a circuit is that the amount of magnetic induction through that circuit should be changed. Now if the magnetic induction is produced by means of a current (either simply, or with the aid of a soft iron core, making it an electromagnet), breaking that current will destroy the magnetic induction as effectually as removing to a great distance. Accordingly making and breaking the current in one circuit (called the primary) will induce currents in a neighbouring circuit (called



the secondary). It is on this principle that the induction coil is constructed: the usual pattern for physiological work is shown in Fig. 87. The primary coil *p*, usually of a few hundred turns, is wound on a wooden bobbin, and is fitted with a core of soft iron wires. The ends of the coil are brought to a pair of terminals *p*. The secondary *s*, consisting of 5,000 or more turns of thinner wire, is wound on a larger bobbin, resting on a foot, which can slide along the baseboard; the ends of the secondary coil are brought to the terminals *s*. The central cavity of the secondary is large enough to allow of its being pushed completely over the primary. When a current flows through the primary it magnetizes the iron wires, and produces a magnetic field of the disposition indicated in Fig. 79.

If the secondary is pushed home, it catches all the magnetic induction proceeding from the primary, but if drawn back along

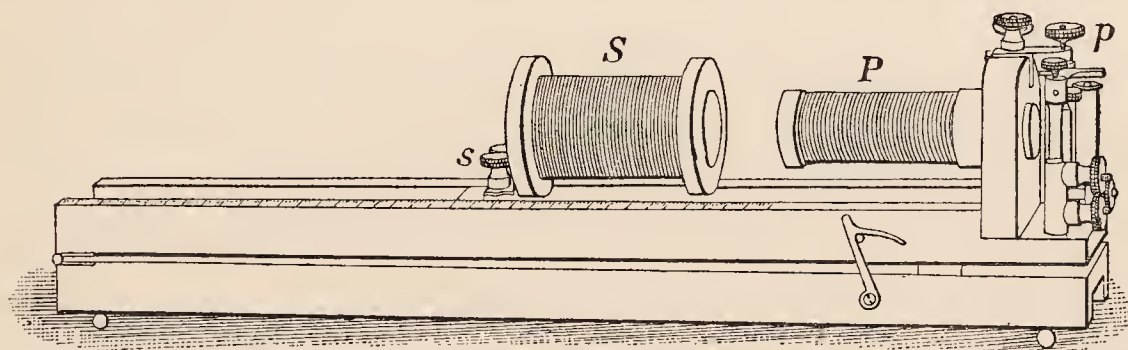


FIG. 87.

the slide, a smaller and smaller fraction of it. This affords a ready means of adjusting the strength of the shocks (i. e. of the electromotive force) in the secondary. The slide is graduated in order to allow of regulating the strength accurately; but it must be remembered that the strengths are not proportional to the distances on it—those must be taken as giving an entirely empirical scale of strength: and further that in order to obtain always the same shock when the secondary is in the same position it is necessary that the current in the primary should be the same. This may best be secured if an accumulator be used to supply current to the primary, but a bichromate, Grove, or Bunsen cell will do, if it be kept in good condition.

If now the battery be connected to the terminals of the primary coil through a simple make and break key, single shocks may be given, by hand. A pair of wires may be taken from the

terminals of the secondary either to a galvanometer, to demonstrate the nature of the induced currents, or to a nerve-muscle preparation. The direction of the induced currents may be found by means of an appropriate modification of Lenz's law; they tend to oppose the change which produces them. Thus when contact is made and the primary current starts, a certain amount of magnetic induction is thrown into the secondary: accordingly a current is induced in the secondary circulating in the sense opposite to that in the primary, for such a current would cause magnetic induction in the opposite sense; it therefore opposes and retards the establishment of the magnetic induction by the primary. It must be remembered, however, that the current induced in the secondary is due exclusively to changes in the magnetic induction through it: consequently when the magnetic field of the primary is fully established the secondary current stops; this may take a few thousandths of a second. When contact is broken the magnetic induction through the secondary is destroyed, a current is induced in the secondary tending to keep up the magnetic induction, i. e. in the same direction as the primary current.

The intensity of the induced electromotive force (and consequently current) in the secondary depends on the rate at which the magnetic induction through it is changing. On making the primary circuit, the increase in magnetic induction is only moderately rapid; but on break the corresponding decrease is extremely rapid, for in a very small fraction of a second the metallic connexion is broken altogether, and the primary current necessarily at an end. Hence, although the same quantity of electricity flows through the circuit at make and at break, it flows much more rapidly in the latter case, and so produces a more violent shock. The effect of the abrupt stoppage of the current is well illustrated by that of abruptly closing a water tap, when the energy of the current of water spends itself in giving a violent jar to the pipe.

When it is desired that the break shock should be of about the same intensity as the make, a modified arrangement due to Helmholtz may be adopted: this is to short-circuit the primary coil instead of breaking the current altogether. The connexions for doing this are shown in Fig. 88. When the key is open the current from the battery can only flow through the coil *pc*, but

when the key is closed, it acts as a short circuit, and diverts practically all the current from the coil; thus the magnetic induction produced by the coil is destroyed, but as the circuit is not broken, the current can die away gradually just as it rises gradually, and the abrupt change usually accompanying the break is avoided. If an accumulator is used to drive the coil, Helmholtz's arrangement should not be used without a resistance (say 1 ohm) in the battery circuit, otherwise the accumulator will suffer from the excessive current taken out of it on short-circuiting.

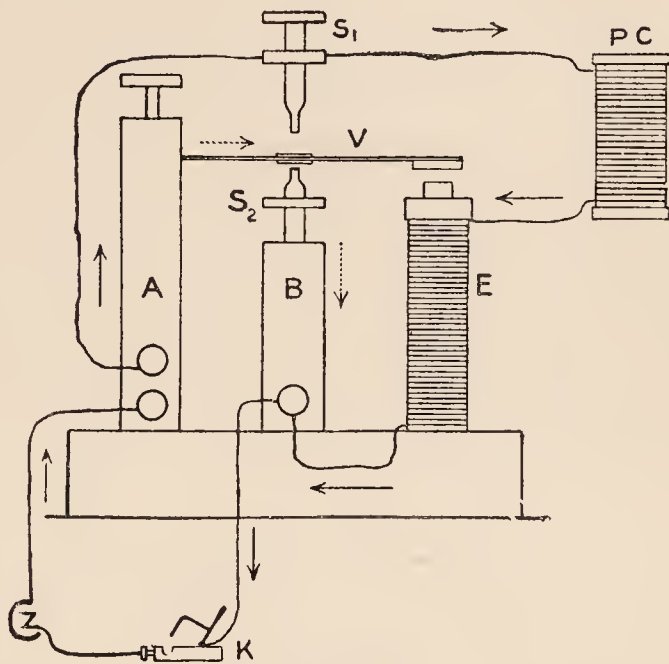


FIG. 88.

It is often desirable to have a rapid series of induction shocks.

This is accomplished by an automatic interrupter, analogous to that of the electromagnetic tuning-fork (p. 4). This sometimes forms a separate piece of apparatus, such as Neef's hammer, shown in Fig. 89. The primary current is led in by a terminal on the brass pillar A which carries a steel spring, goes thence through the screw  $s_1$  to the induction coil PC;  $s_1$  is screwed down till it just touches the spring, and to ensure the contacts against rusting they are made of platinum. From the coil, the current passes round a small electromagnet E and thence by the terminal back to the battery. The electromagnet is arranged so as to attract the

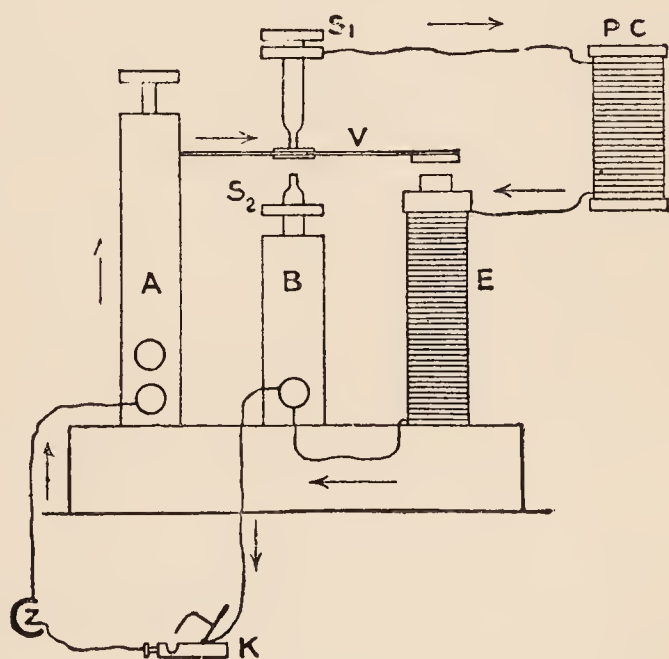


FIG. 89.

spring, and to ensure the contacts against rusting they are made of platinum. From the coil, the current passes round a small electromagnet E and thence by the terminal back to the battery. The electromagnet is arranged so as to attract the



armature of the spring *v*. It does so as soon as the current passes, but in doing so it breaks the circuit and the current stops; the electromagnet then ceases to act, the spring rises again, and the circuit is remade. Thus make and break succeed each other automatically, with a rapidity that depends only on the frequency of vibration of the spring. For Helmholtz's modification it is only necessary to raise the screw  $s_1$  out of the way, and raise the screw  $s_2$  so as to make the vibrating contact through it; and at the same time make a permanent wire connexion from  $s_1$  to the battery as shown. Instead of an independent electromagnet the iron armature of the induction coil itself may be used to actuate an interrupter.

### § 5. Alternating currents.

If an earth coil is rotated at a uniform rate, the magnetic induction through it will alternately increase and decrease, coming back every revolution to the same value. Consequently currents will be induced in it, first one way, then the other, in the course of one revolution, and so on indefinitely. Such currents are called alternating; they consist in a surging to and fro of electricity, instead of a continuous flow in one direction. In the particular case considered, the alternations, or oscillations, of the electricity would be of the kind known as simple harmonic

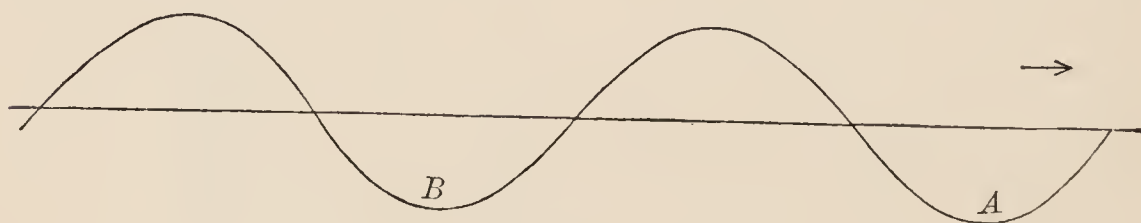


FIG. 89 a (i).

(p. 33). Fig. 89 a (i) may serve as a geometrical representation of an alternating current, the horizontal distances being times, the vertical current strengths, the current in one sense being called positive, in the other negative.

The earth coil is essentially similar to a single armature segment in a dynamo, although, to increase the efficiency, a magnetic field far stronger than that of the earth is used in the latter. Hence in each segment of the armature an alternating current naturally arises. If the ends of the segment are led to two separate metal rings, and brushes be allowed to rest on these,

so that they make a conducting connexion with the armature throughout the rotation, the alternating current may be led away to any external circuit: it is in this way that an *alternating dynamo* is constructed. Current from such dynamos is supplied in many places for house lighting, instead of direct current: the frequency is commonly 100 complete alternations or cycles per second. In a dynamo designed to supply continuous current, instead of the collector rings, it is necessary to provide a commutator, as mentioned in the last section, the function of which is to rectify the current, i. e. to automatically reverse the connexion between the armature and the external circuit every half period. The result of such a commutation on a simple harmonic current is indicated by Fig. 89 a (ii), where alternate halves of the S. H. curves have been inverted. This makes the current flow always in the same sense, although it is far from uniform. In a practical direct-current dynamo there are many sections of the armature, the effects of which are superposed in such a way that when one is giving its maximum effect, another is giving the minimum, and so the total current taken from the brushes is approximately uniform in strength.

If alternating current be supplied to the primary of an induction coil, an interrupter is needless, for, as the magnetic induction due to that cur-

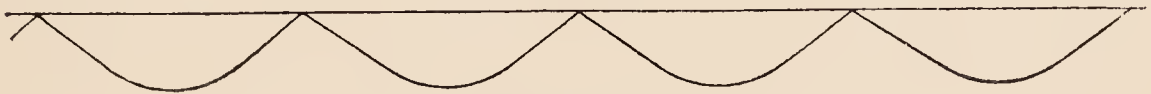


FIG. 89 a (ii).

rent will be constantly changing, and will pass through the secondary, alternating currents will be induced in the secondary too. Here there is no 'make' and 'break,' but a continuous rise and fall: such induced currents are not, therefore, so effective in stimulation as the more abrupt ones produced by the ordinary method; but the principle involved is extensively used for another purpose. The *transformer* is essentially an induction coil without an interrupter; but the primary and secondary are interwound, in order to make sure that there is no 'magnetic leakage,' i. e. that all the magnetic induction proceeding from the one flows through the other—a condition only obtained in an induction coil of the pattern described above when the secondary is pushed home. Transformers are used to convert low voltages into high, and vice versa: approximately, the E. M. F.'s in the primary and secondary circuits are in the same ratio as the number of turns of wire in them. It is economical in practice to generate alternating current, for domestic use, at a high voltage—say 2,000—and then, as it would be dangerous to introduce so high a voltage into houses, transform it down to, say, 200 volts, by a transformer placed near to each house supplied. In this case, the ratio of transformation being 10:1, there must be ten times as many turns in

the primary coil as in the secondary—a state of things the converse of that in the ordinary induction coil, where the secondary has by far the larger number, and consequently the instrument transforms ‘up.’

It must be remembered, however, that a transformer, like any other machine, is not capable of giving out more power than is put into it; hence the power, measured by the product of the volts into the ampères, must be (ignoring the imperfections of the apparatus) the same on the primary and on the secondary side, whether the appliance is for transforming up or down. Thus, if a transformer receive on the primary side 2 ampères at 2,000 volts, i. e.  $2 \times 2000 = 4000$  watts, and the ratio of transformation be 10 : 1, it should give off the secondary 20 ampères at 200 volts.

A transformer is an obviously reversible appliance; either side may be made the primary, at will. Hence the same transformer may be used to transform up or down.

It has already been mentioned that alternating currents are required in order to measure the resistance of electrolytes (p. 221), the arrangement being a Wheatstone bridge, in which the secondary of an induction coil is substituted for the battery, and a telephone for the galvanometer. The coil used for this purpose should be an extremely small one, capable of being driven by a single Leclanché cell, and should have a very high rate of interruption, 500 a second or thereabouts, as that is about the pitch of the human voice, for which telephones are adapted: this is secured by having a very light contact spring *v*, Fig. 88, not loaded. The core of the coil itself is used as electromagnet.



## CHAPTER IX.

### ELECTROSTATICS.

#### § 1. Forces between electric charges.

IF the poles of a battery be connected by wires to conductors of any shape—say a pair of parallel metal plates—near to one another, but not touching, of course no continuous current will flow, since the circuit is broken by an air-space ; but a momentary current will flow, producing a *charge* on the plates, i.e. a quantity of positive electricity on one, an equal quantity of negative on the other. An *electrostatic stress* is set up in the air-space between the plates, and, when the amount of this is sufficient to balance the electromotive force of the battery, the action stops. The process may be illustrated by the parallel of a circuit made up of materials of varying conductivity: suppose an accumulator connected by thick copper wires to an incandescent lamp: then the electromotive force of the accumulator is distributed in proportion to the resistance to be overcome, and, as that of the lamp is far greater than that of the wires, it is practically all thrown on to the lamp. Now for the lamp substitute an air-space, as above: then all the electromotive force is spent on the air-space, or in other words there is a difference of potential between the plates equal to that of the battery. This difference of potential produces a strain in the space<sup>1</sup> between, which shows itself in the phenomenon of a charge on the plates. Just as in the magnetic case we had to consider the existence of a field of magnetic force between magnet poles, and were able to trace it out by means of lines of magnetic force, so here there is a *field*

<sup>1</sup> It is not the air that is strained, for the phenomena are just as well observed in a vacuum: it is rather the 'ether,' the medium that is assumed to fill all space, and transmit light as well as electrical actions.

of *electric force* (or stress) between the charged plates, and we can trace it out by means of lines of force leading from the one to the other. These lines indicate the direction in which the strain lies, and necessarily reach across from the one conductor to the other, because a *conductor is a material that cannot support an electric stress*; consequently where the line of force reaches a conductor it ends—the state of strain ceases—and it follows also that each positive charge of electricity must be associated with an equal negative charge, for that is only to say that one cannot have one end to a line without having the other end.

If the conductors on which the charge is located have any definite size and shape, the quantity of electricity stored up

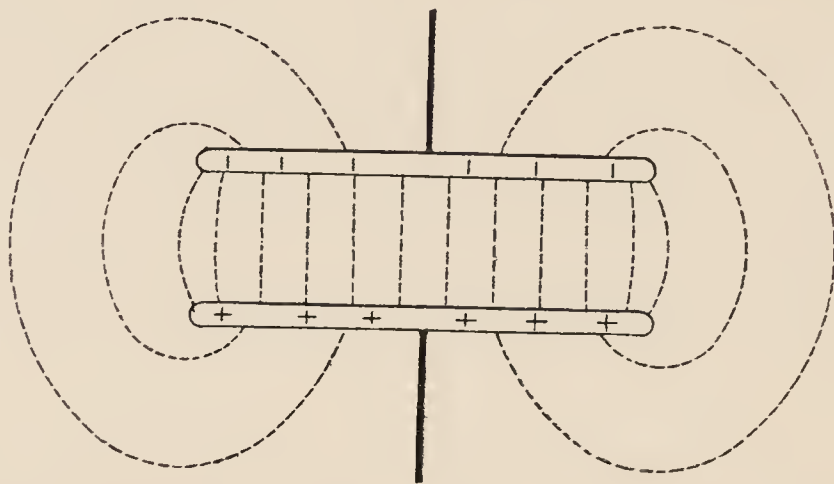


FIG. 90.

on them will be proportional to the difference of potential applied, and the ratio  $\frac{\text{charge}}{\text{difference of potential}}$  is called the *capacity* of those conductors. When the conductors are arranged so as to have a large capacity, they form what is called a *condenser*; this is best attained by making the metallic surfaces large, and the air-space between them small, the capacity being proportional to the area, and inversely to the thickness of the layer of non-conductor.

The stress that exists in the non-conductor, or *dielectric* as it is usually called in this connexion, is a tension along the lines of force, combined with a pressure across them. Thus in Fig. 90, which represents the electric field between a pair of

plates, and Fig. 91, for a pair of charged balls, and similar figures, we are enabled by drawing the lines of force<sup>1</sup> to see at a glance what forces will act on the charged bodies. We must imagine that each line of force has a tendency to shorten

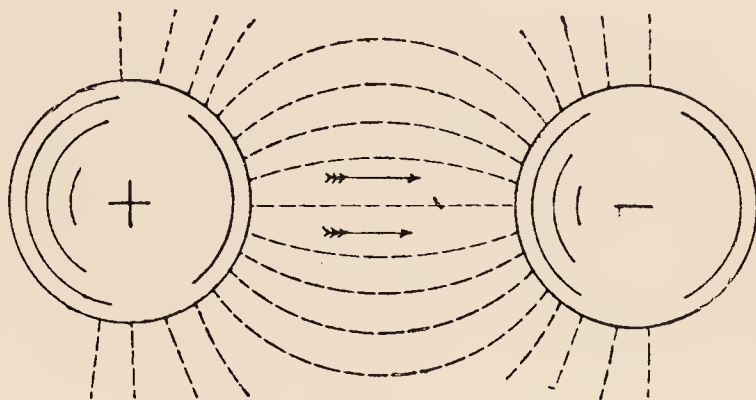


FIG. 91

up, like an elastic thread, but that adjacent lines push each other apart sideways. It is evident that the positive and negative charges on the plates or balls attract each other. On the other hand, if two conductors, both charged with the same kind of electricity (e.g. two balls connected to the same pole of a battery), be placed near together, there is a repulsion between them (Fig. 92).

The precise law of force is identical with that of magnetism, viz. the *force exerted between two charged bodies is equal to the product of their charges divided by the square of the distance between*

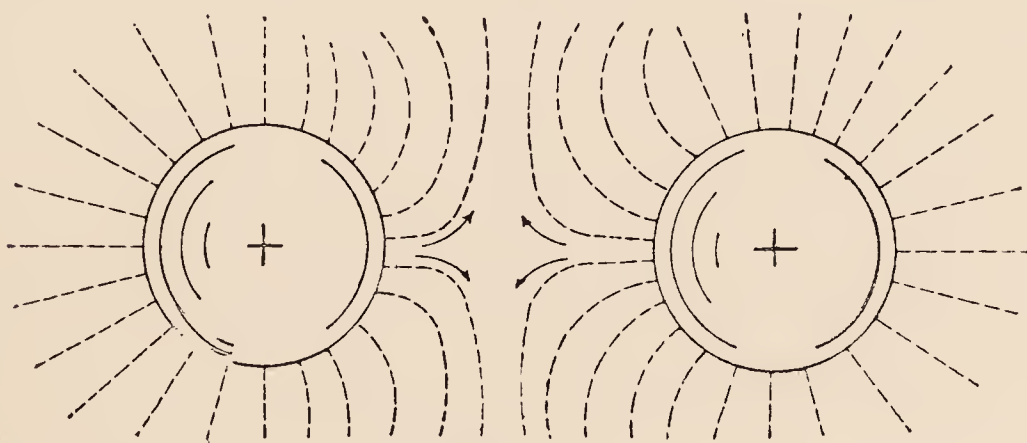


FIG. 92.

*them*: when the charges are of the same sign the force is a repulsion; when of different signs, an attraction. This law

<sup>1</sup> The lines of force are identical in arrangement with those of magnetic forces which would be produced if each quantity of electricity were replaced by an equal quantity of magnetism.



gives a definition of unit quantity of electricity, viz. that quantity which placed 1 cm. from a similar and equal quantity repels it with a force of 1 dyne. This is the C. G. S. electrostatic unit of quantity: that which we have previously used is either the electromagnetic unit of the C. G. S. system or the practical unit, the coulomb, which is  $\frac{1}{10}$  of the electromagnetic unit. The electrostatic unit is far smaller—30,000,000,000 times smaller—than the electromagnetic unit. The quantities of electricity practically occurring in electrostatic experiments, e.g. in charging a condenser, are in fact excessively small compared with those that flow in currents of ordinary magnitude.

On the other hand, the differences of potential required for electrostatic experiments, such as demonstrating the attractions and repulsions of charged bodies, are usually great—hundreds or thousands of volts. The connexion between the potential difference applied and the state of electric force or stress set up in a non-conducting space may be most conveniently illustrated by means of the parallel plate condenser already referred to. It must be borne in mind that difference of potential or electromotive force (see the definition, p. 193) means the work done in carrying unit quantity of electricity from one point to another: if now the plates be put at a fixed distance apart and a fixed E. M. F. be applied to them, an electric force, or stress, is produced in the intervening space measured by the applied E. M. F.  $\div$  distance apart. But the work done in moving from one point to another  $\div$  distance is the measure of the mechanical force overcome. Hence, just as E. M. F. is the expression of the work done in transporting a unit of electricity, so the electric force (or stress, or intensity, or strength of field) is the expression of the mechanical force exerted on a unit of electricity. It is therefore a quantity precisely comparable with the magnetic force, which is the mechanical force exerted at a point on a unit of magnetism. Now, suppose the distance between the plates to be doubled while the applied E. M. F. is kept the same: then the electric force between the plates is halved. But, as we have seen, the 'charges' on the plates are merely the ends of the lines of force; if the force is halved in intensity, that means that there are only half as many lines of force, and the charge must be halved in amount. Thus we arrive at the result that the charge on a parallel plate condenser, to which a given E. M. F. is applied,

varies inversely as its thickness. But the charge for unit E. M. F. is the quantity we have defined as the capacity of the condenser. So we may say that the capacity varies inversely as the thickness.

Capacity in electrostatics plays a part analogous to that of conductance in current electricity: for an electromotive force applied to a dielectric produces a charge, applied to a conductor produces a current; and the ratio of charge (resp. current) to E. M. F. measures the capacity (resp. conductance). Hence all the rules given previously for the combination of conductances are applicable to the combination of capacities.

The main difference between the laws of electric charges and magnetism is that whilst on breaking a magnet each fragment becomes a complete magnet, so that it is impossible to separate the north and south poles, that is not true in the electrical case. If a pair of plates be charged by means of a battery and the plates then insulated from the poles of the battery, by removing or breaking the connecting wires, the plates will remain charged, and we shall have an isolated positive and an isolated negative charge, which may be carried about independently.

It is found that merely rubbing two bodies together sets up an electromotive force between them, so that on separating the bodies they are found to be charged. And as the electromotive force thus produced is exceedingly large (thousands of volts), although the quantity of electricity generated is very small, this process is most conveniently adapted for performing electrostatic experiments. A glass or ebonite rod rubbed with a catskin gives good results, provided both are sedulously dried beforehand; the glass or ebonite becomes negative, the fur positive. It is therefore convenient, in demonstrating the forces between electric charges, properties of condensers, &c., to use such a generator of electricity, instead of the enormously large voltaic battery that would be needed.

Another important means of obtaining electric charges is by what is known as *electrostatic induction* (not to be confused with induction of currents, described above). If a charged body, say a metal knob, supported by an insulating stand, charged positively by means of a glass rod and catskin, be placed near another conductor, it will attract, or induce negative electricity on the part of the other conductor nearest to it, and drive away



an equal amount of positive to the far end of the second conductor. If the latter be insulated, these two charges remain in equilibrium so long as the original charge is in place, but if that be removed, they reunite, and the second conductor is again in the neutral condition. But if while the original or inducing charge is still there, the conductor be connected to the earth, say by touching it with the finger, the positive induced charge will be driven further away, now that a conducting path is open to it, on to the earth; then if the finger be removed the negative induced charge will be left by itself on the second conductor, and when the inducing charge is taken away, the second conductor will remain negatively charged.

The simplest machine constructed on this principle is the *electrophorus*. This consists of an ebonite disc, usually provided with a metal backing; and a metal disc of the same size, fixed on a long glass handle. The ebonite disc is laid on a table and rubbed with fur or silk to give it a negative charge. The metal disc, held by the insulating handle, is then placed on it: the charge does not leak on to the metal quickly because ebonite is an exceedingly bad conductor, and also there is usually contact between the ebonite and metal only at a few points, and consequently a thin film of air between the two over the greater part of the surface. The action consequently is to induce positive electricity on the lower, negative on the upper surface of the metal. Now touch the upper surface for a moment with the finger: the negative charge flows away, leaving only positive electricity on the metal plate. If the metal plate be now lifted up by the insulating handle so as to avoid leakage, it will be found to be charged; moreover, as the metal and ebonite plates, being very close together, form a condenser of large capacity, there will be a somewhat large charge, and on separating them the capacity is much reduced, so that the potential is correspondingly increased. The negative charge remains on the ebonite, and can be used over and over again, till it disappears through leakage.

The more elaborate electrical machines, such as those of Holtz and Wimshurst, are developments of the principles used in the electrophorus. In making electrostatic experiments an *electroscope* is made use of, i. e. an instrument for observing the potential of charged bodies. The simplest and most useful



pattern is the gold-leaf electroscope; this consists essentially of a pair of gold leaves hung from a metallic support, which in turn is insulated. It may easily be made as follows:—a conical flask of good insulating glass (shellacked to keep the surface dry, and so improve the insulation further) is fitted with a rubber stopper; through this passes a stiff brass wire; a brass knob or plate is soldered to the top of the wire, while below it is bent so that a pair of strips of gold-leaf may conveniently be attached to it, and allowed to hang down side by side. If now the knob be connected to a source of high potential, positive or negative, or charged by means of a rubbed ebonite rod, a charge will be imparted to each gold-leaf, and these repelling one another, the leaves will stand apart at an angle. Gold-leaf is chosen on account of its extreme lightness.

Such an electroscope may be provided with a scale, by which to measure roughly the potential to which the leaves are raised: it then becomes an *electrometer*.

But a better instrument is that known as the 'quadrant.' This, as usually constructed, consists of four quadrants, seen in plan in Fig. 93, forming a nearly closed cylindrical box, but with air-spaces between: centrally to them, the 'needle,' a flat sheet of aluminium shown by the dotted line, is sus-

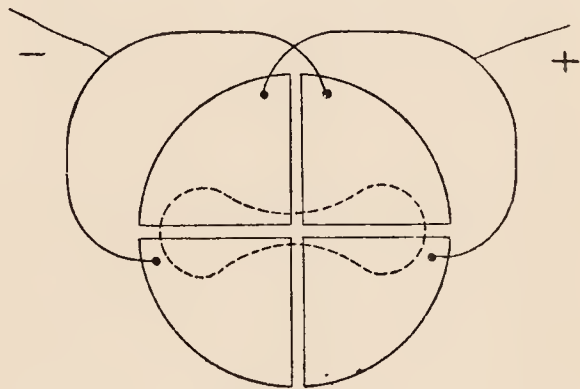


FIG. 93.

suspended, so as to lie between the upper and under surfaces of the box. The quadrants are rigidly supported and insulated: one pair, diagonally opposite to each other, are to be connected to the positive, the other pair to the negative end of the source of electromotive force to be measured. The needle is supported by two parallel silk fibres, which serve at the same time to insulate it. A wire attached to its lower side dips in a dish of strong sulphuric acid; this serves three purposes, as a means of conveying a charge to the needle, as a dash-pot to prevent oscillations of the needle, and as a desiccator to maintain good insulation in the instrument. The needle is charged by a rubbed ebonite rod or otherwise, and the quadrants connected, as shown: then the (negative) charge on the needle

is attracted by the two positive quadrants, repelled by the two negative, with the result of giving a small rotation to the needle. This is measured by mirror and scale as in a galvanometer. The sensitiveness is usually about 100 scale divisions per volt.

We have so far tacitly assumed that the dielectric, in which the phenomenon of electric stress shows itself, is air, or if not so a vacuum. There are, however, many solids and liquids that can be used instead, and which are comparable, as non-conductors, with gases. The chief difference produced by substituting one of these for air, is that the strain corresponding to a given stress is increased; this is analogous to the differences that exist between different solids in their ordinary elastic properties. Up to the present we have identified stress and strain in the electric case, and spoken indifferently of the stress or electric force, as measured by the difference of potential per unit length in the dielectric, and the effect it produces, or strain, showing itself in the form of a charge on the conductors bounding the dielectric. If now we have to do with solids or liquids we must not overlook this distinction. A given stress (volts per cm.) produces a greater strain or charge (coulombs per sq. cm.), and the ratio of the charge produced to that which would be produced if there were air present is called the *dielectric constant*. Since  $\text{charge} \div \text{potential}$  is defined as capacity, we may express the same idea by saying that the capacity of a condenser is greater when it has a solid or liquid dielectric than when there is air between the plates, and that the ratio of the capacities is called the dielectric constant: for this reason the term specific inductive capacity has been given to the same quantity. Thus, e. g., a condenser formed of two brass plates in air has a certain capacity: if the space between the plates be filled with benzene the capacity is about 2.25 times greater. A very common form of condenser is the Leyden jar, which is a glass cylinder, coated inside and out with tin-foil: the glass here forms the dielectric, and its dielectric constant is from 6 to 8.

## § 2. Electric discharges.

There is a limit to the electric stress that any material can stand. This limit for ordinary air may be put roughly at 30,000

volts per cm. ; if this be exceeded, a discharge, usually in the form of a spark, takes place. Thus the sparks given by an induction coil when the wires leading from the secondary are held near together, show the discharge current, and the spark length is roughly a measure of the E. M. F. that the coil can give.

If a Leyden jar be charged, and the two coatings be touched by a wire, of course a current flows through the wire, and the jar is discharged. But the sudden rush of electricity that takes place when a conducting channel is provided, behaves like a pendulum pulled aside and then released : it does not merely settle down to the equilibrium state, but overpasses it and swings to and fro. Thus, there is an alternating current in the wire, which only gradually dies away, leaving the two coatings of the jar at the same potential. This at least is what happens if there is not too much resistance offered to the discharge ; but if the connecting wire were very long and thin, the current might be so much damped that no oscillations would take place ; just as a pendulum might be so damped, by immersing the bob in glycerine, that it would only slowly settle down to the equilibrium position without swinging past it. The discharge from a Leyden jar may take from  $\frac{1}{10000}$  to  $\frac{1}{1000000}$  of a second to execute one complete oscillation : in any case the frequency is far greater than that of the alternating currents derived from an ordinary dynamo.

But far more rapid alternations than this were obtained by Hertz. He connected the terminals of an induction coil to a condenser, consisting of two large metal balls ; these were placed a metre or so apart, and a pair of thick wires arranged so as nearly to connect them, leaving however a small air-gap in the middle (Fig. 94). The action of the coil

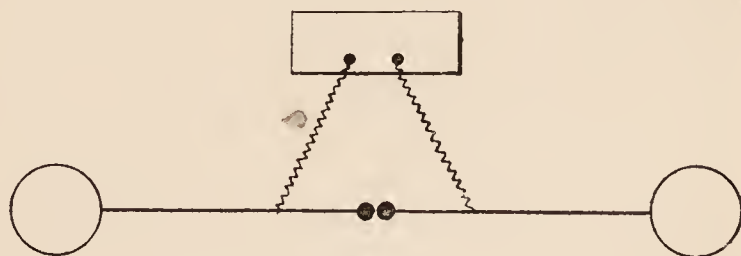


FIG. 94.

charged the balls up to a high difference of potential, and then a spark crossed the gap : but now it has been observed that though air is normally a perfect non-conductor, when a spark is passed through it it becomes for a short time a fairly good



conductor. Consequently, when the resistance of the air-gap was once broken down, the electric current set up could surge backwards and forwards between the balls a good many times before it was frittered away. The frequency of the oscillations thus produced was about a hundred million per second, or even more.

Such oscillations set up waves in the surrounding medium (ether), just as a vibrating tuning-fork sets up waves of sound in air. Hertz was able to detect these waves, investigate their properties, and in particular to show that they are propagated with a velocity of  $3 \times 10^{10}$  cms. per sec. This is the velocity of light, and in fact the electromagnetic radiation turned out to have all the properties of light, except that it was on an enormously greater scale. Hence we may conclude that light is really an electromagnetic radiation, set up by the vibrations of molecules, which play a part similar to that of Hertz's apparatus: and as the molecules are very much smaller, they send out waves that are shorter, and consequently of far higher frequency even than the Hertzian waves.

When air is reduced in pressure, its resistance to electric stress is reduced too; therefore *vacuum tubes* are made, with which many phenomena of discharge can be obtained which are not possible in air at atmospheric pressure. Such a tube consists of a glass vessel, provided with electrodes of platinum or aluminium fused through the glass; it is then evacuated by a mercury pump, and hermetically sealed. Air in small traces that the pump will not remove may be left in it, or the air may be replaced by traces of some other gas or vapour. When an induction coil is connected to the electrodes of such a tube, discharges can be produced through it, varying very much in shape and character, according to circumstances, but usually attended by an appreciable amount of light. Of all the numerous phenomena that have been observed in this way, we can only refer to one—the production of Röntgen rays. If a very good vacuum be produced in the tube, the negative electrode (called, as in electrolytic cells, the cathode) sends out a discharge which proceeds in straight lines, as if consisting of fine particles shot out from the electrode with a high velocity; these, striking the glass of the tube, produce on it a characteristic green phosphorescence, and there generate Röntgen rays, which also travel

in straight lines, and penetrate all substances to an extent inversely proportional to their density. Thus bone, being comparatively dense, obstructs the rays and produces a shadow, which the lighter tissues do in much less degree, while a metal object, such as a needle inside the hand, throws a deep shadow to the Röntgen rays.

An ordinary induction coil can be used to excite the vacuum tube, but one of much larger size and much stronger than the ordinary physiological pattern is desirable. But if used with an ordinary spring interruptor, the number of discharges given by the coil per second makes the process of photographing by Röntgen rays slow. Wehnelt's electrolytic interruptor has therefore been generally adopted instead. This consists of a pair of electrodes in a jar of dilute sulphuric acid: the one, of lead, is of large size, the other, platinum, is very small, being only a few millimetres of thick wire projecting into the liquid. When a current passes, the small electrode polarizes far more rapidly than the big one, and a bubble of gas forming there, stops the current; this escapes, and the current is renewed. By putting such an interruptor in the primary of the induction coil, discharges at the rate of 1,000 and more per second are obtained, and the vacuum tube for generating Röntgen rays becomes proportionately more effective.

## CHAPTER X.

### LIGHT.

#### § 1. Production of light.

WE saw, at the end of the last chapter, that it is possible to produce waves of electromagnetic disturbance, which possess all the characters of ordinary radiation, except that they are of very much less frequency, and consequently greater wavelength. Light, or more generally speaking radiation, is now known to consist of electromagnetic waves, set up by vibrations of the molecules, and transmitted through the all-pervading medium known as the ether. Since light consists of waves, much that was stated of sound is applicable here also, and it is well to keep the statements of that chapter in view.

In the first place, then, the vibrations producing light can differ in (i) amplitude, (ii) frequency. As in sound, the energy or intensity of a vibration depends on the amplitude, being proportional to the square of it. It is a matter of common observation, that by raising the temperature of any body it may be made to emit light, and the hotter it is the more light it emits; thus an incandescent lamp is made to glow by running a current through it, so as to make it hot: if the current is increased in strength, the lamp becomes hotter, and gives out more light. This means, then, that the vibrations executed by the molecules of the lamp filament become greater in amplitude and so possess more energy, enabling the molecules to radiate energy in the form of light more vigorously.

Varying frequency of vibration, which in sound produces the sensation of pitch, in light produces that of colour. If a vacuum tube be made with a trace of hydrogen in it, and an electric discharge be passed through it, it glows, and if examined by means of a spectroscope (vid. inf.), it will be found that it is



giving out three distinct kinds of light, one red, one blue, one violet, just as a bell when struck might give out two or three partial tones of different pitch. Of the three, the red has the lowest frequency, violet the highest. It is however customary to speak of the wave-length rather than the frequency in the case of light: and it must be remembered that the wave-length = velocity of wave  $\div$  frequency, consequently the red has the longest wave-length, violet the shortest. The actual wave-lengths are—

Hydrogen, red (C)	$6.563 \times 10^{-5}$ cm.
„ blue (F)	4.861
„ violet	4.341

(The letters C and F are names by which the particular kinds of light are sometimes known, see p. 276.) It is not, however, generally true in light as in sound, that a substance gives out only one or a few vibrations of definite frequency. When a solid is heated to whiteness, the radiation from it is found to be of all possible wave-lengths between certain limits; it is as if one played all the notes of a piano at once. The range of wave-length observed is from  $7.6 \times 10^{-5}$  cm. (deepest red), through orange, yellow, green, blue, violet, to about  $3.8 \times 10^{-5}$ . This series of colours is called the spectrum, and we may speak of going ‘up’ the spectrum from red to violet, by analogy with going up the piano from notes of low frequency to those of high. But the limits just mentioned are really in no way inherent in the physical phenomenon of radiation: they are purely physiological. Just as the ear is incapable of appreciating aërial waves above and below certain limits of frequency, so with the perception of ethereal waves by the eye. Waves of length greater than  $7.6$  and less than  $3.8$  exist, and can be observed by appropriate instruments, consequently the spectrum really extends beyond the visible portion, and includes an ‘infra-red’ region (of long wave-length), and an ‘ultra-violet’ (of short wave-length).

When a solid is heated, it first gives out radiation only of very great wave-length; as the temperature rises the spectrum extends further up, coming to include shorter waves, till at between  $400^{\circ}$  and  $500^{\circ}$  cent. the limit of visibility is reached, and the solid appears a dull red. As the temperature is further raised, other colours of shorter wave-length are produced, and the body appears, as a consequence of their blending, first yellow,

then white, and, even in an extreme case, slightly bluish, owing to preponderance of the blue and violet in the spectrum. It is only at very high temperatures that an appreciable amount of ultra-violet is produced.

## § 2. Light-waves.

When radiation proceeds from some source, such as an incandescent solid, it may sometimes be made to travel in one direction only, and not spread out. That may be accomplished by a properly placed concave reflector, or by a lens, as is used in light-houses; it also occurs when light is generated at one end of a glass rod, especially if the rod be silvered: the light is led along the length of the rod, with little loss at the sides, even if it be somewhat bent—a contrivance sometimes used in microscopy and in surgery. In either case, since the light does not spread out, it remains of the same intensity, however far it travels, except in so far as the substance conveying it absorbs some of the energy of the waves, i. e. is not perfectly transparent. But unless some precaution is taken, light spreads out in all directions, so that the energy of the waves has to be spread out over larger and larger surfaces, as the light travels further from its origin, and becomes attenuated. For the same reason as in sound, the intensity of the light is inversely as the square of the distance of the source, disregarding any further diminution of intensity due to want of perfect transparency.

This fact is made use of in order to measure the strength, or ‘candle-power’ of lamps. Any apparatus for this purpose is called a *photometer*. It is not possible to estimate directly by eye the ratio between the intensities of illumination produced by two lights, but it is possible to tell when the two illuminations are the same. Hence, in a photometer, the brighter light is taken gradually further away, till it produces the same intensity of illumination as the other: the candle-powers of the two lights are then in proportion to the squares of their distances. The most familiar arrangement for comparison is the Bunsen photometer: this consists of a sheet of white paper with a spot of grease in the middle. When illuminated from the front the grease spot looks darker than



the rest of the paper, when from behind it looks lighter; if the illumination back and front be equal, the grease spot is hardly visible. Now suppose a standard candle to be fixed 50 cm. from one side of the paper, and an incandescent lamp moved to and fro on a graduated bar, extending on the opposite side of the paper, and suppose that at a distance of 200 cm. the grease spot disappears; then the ratio of distances is 4 : 1, and the candle-power of the lamp is  $4^2 = 16$ .

The velocity with which light is propagated is enormously great: it is found to be the same for light of all wave-lengths (including artificial electromagnetic waves) when traversing empty space. This common velocity is about  $2.999 \times 10^{10} \frac{\text{cms.}}{\text{sec.}}$ .

Remembering then the relation between velocity, wave-length, and frequency, we are able to calculate the latter. In the case of red hydrogen light it is  $2.999 \times 10^{10} \div 6.563 \times 10^{-5} = 4.57 \times 10^{14}$  per second: this is, of course, far beyond anything that can be produced artificially; it is, in fact, about a hundred thousand times more rapid than the most rapid Hertzian waves that have been produced.

In media other than empty space, the velocity of light is less, and moreover it varies slightly according to the kind of light considered. These relations are most conveniently expressed by means of what is known as the *refractive index* of the medium; this may be defined as the ratio  $\frac{\text{velocity of light in vacuo}}{\text{velocity in the medium considered}}$ .

The refractive index for any given substance, say water, varies for light of varying wave-length: the general rule is that it is greater for shorter waves, i.e. greater for blue than for red. The variation is called the *dispersion* of the medium, since it is the means practically adopted for separating or dispersing light of different kinds. As the refractive index and dispersion play a most important part in the construction of optical instruments, we subjoin a short table of numerical values. It will be noticed that the refractive index of air is so nearly unity that it is rarely necessary to take into account either it or the dispersion of air, and the same is true of other gases; so the refractive indices of solids and liquids are often measured by comparison with air instead of with a vacuum.



	Refractive index for sodium light.	Dispersion from hy- drogen blue to red.	$\frac{\mu_D - 1}{\mu_F - \mu_C}$
	$\mu_D$	$\mu_F - \mu_C$	$\mu_F - \mu_C$
Air at 0° and 760 mm.	1.000291	0.00000288	101.
Water at 20°	1.3330	0.00600	55.5
Silicate crown glass (Schott & Co. O 40)	1.5166	0.00849	60.9
Ordinary silicate flint glass (Schott & Co. O 118)	1.6129	0.01660	36.9
Carbon disulphide at 20°	1.628	0.0342	18.4

The second column gives the refractive index for the yellow light of incandescent sodium vapour of a wave-length intermediate between hydrogen blue and red.

The third column gives the difference between the refractive indices for blue and red light; the fourth column gives the ratio (mean refractive index - 1)  $\div$  difference for blue and red; the usefulness of this will appear later.

A difficulty was at first experienced in accepting the view that sound and light are both due to wave motions, in the fact that sound travels round corners with very little difficulty, whereas light, when it meets an opaque obstacle, casts a sharp shadow. The difference, however, is not a real objection; it is all a matter of scale. Sound-waves are ordinarily a metre or so long, light-waves about  $\frac{1}{20000}$  of a centimetre; hence, to perform comparable experiments in the two cases, it is necessary to make the apparatus enormously large in the one case, or very small in the other. If a church bell be taken as a source of sound, and a house as obstacle, it will be found that quite a marked sound shadow is formed—the sound is much less intense at points where the house stands between the observer and the bell. Still more marked is the shadow thrown by a hill to the sound-waves from a large explosion of gunpowder. If, on the other hand, the shadow thrown by a small sharp-edged object to a source of light be examined minutely, it will be found that a little of the light does get round corners. The experiment may be made as follows: paste tin-foil on a sheet of glass, and cut a narrow slit in the foil with a knife; mount this slit in front of a very brilliant source of light; some 20 or 30 cms. away mount a sharply cut straight edge of metal, taking great care that the edge is parallel to the slit; then examine the shadow behind the edge with a fairly high power eye-piece, holding this so as to look in the direction of the light, and to have half its field of view in the light, half in the shadow. It will be found that there is no sharp line of demarcation between light and dark, but that the light fades away quite gradually towards the dark side, and that on the light side a series of alternations of light and dark occur, getting fainter as one looks further out into the

illuminated part of the field. These alternations are called *diffraction fringes*, and it is found that a complete explanation of them can be given on the theory that light consists of waves, so that, far from being a difficulty, the phenomena of shadow-formation constitute the strongest evidence in favour of that theory.

Accordingly we may say that light produced at a source in the air travels outwards in spherical waves, with uniform velocity. We may trace the course of the light by means of the *wave-front*, i. e. the surface including all the points that are in the same phase at the same moment (see figures, pp. 155 to 157): the wave-front is a sphere round the source of light, but, if the source is far away, and we are only considering a small piece of the wave-front (e. g. the piece of a wave-front from a star, that enters a telescope), we may regard that piece as plane. Further, we may speak of *rays* of light, lines at right angles to the wave-front, and consequently in the direction in which the light is going: we may trace out the course of the light by means of the rays, if more convenient; and, if the wave-front meets an opaque obstacle, that part of it may be considered extinguished, and, very approximately, a shadow will be formed which is marked out by the rays that pass the edge of the obstacle. The space marked out by these rays is called the geometrical shadow, and, as we have seen, light really bends into this space, but only to a minute extent: practically the light may be regarded as travelling in straight lines, i. e. along the rays.

When a wave of light strikes against a surface constituting the boundary between two media it is in general partly reflected, and partly passes on into the second medium. The amount reflected varies very much—if the surface be of polished silver it may be more than 90% of the whole.

The geometry of the reflection is shown in Fig. 95.  $AB$  is the reflecting surface;  $EF$  is a small portion of a wave-front, such as may be regarded as

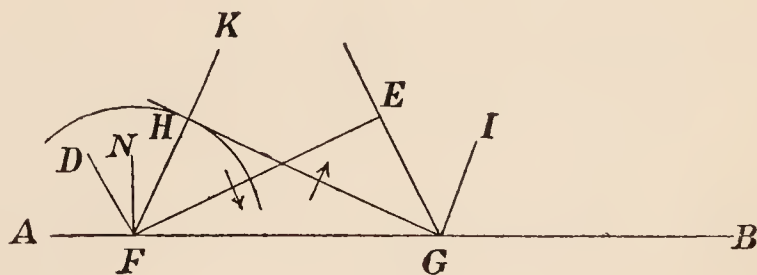


FIG. 95.

plane (the diagram is to be regarded as a section at right angles to the plane of the reflecting surface, or of the wave-

front). EF is travelling in the direction shown by the arrow, and at the point F has reached the reflecting surface; the light from E only does so later, viz. when it has reached the point G. Now from F as centre draw a circle with radius equal to EG, and from G draw the tangent GH to this circle. Then, by the time the light from E has reached G, that from F reflected has necessarily covered the distance FH, and GH is the new wave-front, i. e. surface, all points of which are reached simultaneously by the light. This is travelling in the direction of the arrow, along FHK or GI. Hence, we see (by symmetry) that the reflected wave-front makes the same angle with the reflecting surface that the incident wave does: EFG is called the angle of incidence, HGF the angle of reflection: these two are equal. This result may also be expressed by means of rays; e. g. DF is a ray of the incident light (perpendicular to the incident wave-front), FK a ray of the reflected light. These, then, make equal angles with the normal FN to the surface, and since  $DFN = EFG$ , and  $HFN = HGF$ , we may equally well define the angles of incidence and reflection as the angles which the incident and reflected rays make with the normal to the surface, and state the law of reflection as follows:—*The incident and reflected rays are in the same plane with the normal to the surface, and make equal angles with it, but on opposite sides.*

Light-waves that strike a boundary between two media break up in general into a reflected part and a transmitted part, but, as

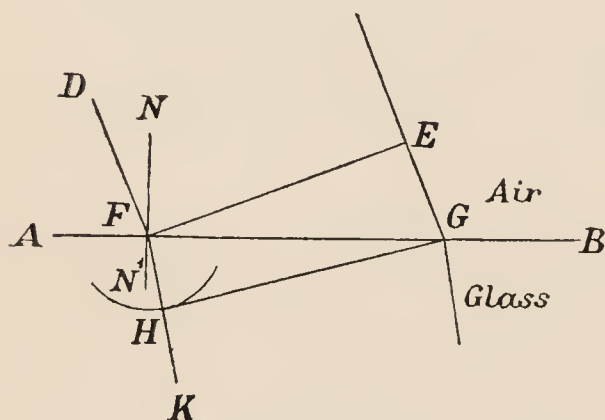


FIG. 96.

many substances are opaque, the latter is often rapidly absorbed and cannot be observed. If, however, the second medium be transparent, e. g. glass, the transmitted beam of light can be observed, and is usually much the stronger of the two; it does not in general proceed straight, but is said to suffer *refraction*. The cause of this

is the unequal velocity of light in the two media, as may be explained with the aid of Fig. 96. Hence, as before, AB is the trace of the bounding surface, EF that of a plane



wave-front falling on it. But during the time that the incident light takes to travel from E to G in air the light from F travelling in the glass will only have gone a shorter distance. Let the refractive index  $\mu = 1.5$ , then the light in the glass will travel only  $\frac{2}{3}$  as fast, and we must draw round E a circle of radius equal to  $\frac{2}{3}$  of EG; then the tangent from G on this circle, viz. GH, is the refracted wave-front. It is evident that this makes a smaller angle with the refracting surface than the incident wave-front does. If we call EFG the angle of incidence,  $\angle_1$  and FGH the angle of refraction  $\angle_2$ , then  $\sin \angle_1 = EG \div FG$ , while  $\sin \angle_2 = FH \div FG$ , so that

$$\frac{\sin \angle_1}{\sin \angle_2} = \frac{EG}{FH} = \mu.$$

Putting this in terms of the rays, we may define  $\angle_1$ , the angle of incidence, as the angle between the incident ray DF and FN normal to the refracting surface, while  $\angle_2$ , the angle of refraction, is the angle between the refracted ray FK and FN<sup>1</sup> the normal, and say:—*The incident and refracted rays are in the same plane with the normal to the surface, and the angles of incidence and refraction are connected by the relation  $\sin \angle_1 = \mu \sin \angle_2$ .* Here, in whichever sense the light travels,  $\angle_1$  is to be taken as the angle in air,  $\angle_2$  in the other medium, as may be seen from the nature of the above proof.

The consequences of the laws of reflection and refraction will be considered in detail below in dealing with optical instruments.

It was mentioned, in dealing with sound, that vibrations in air are necessarily longitudinal. Those of light, on the other hand, are transverse, that is to say, the movements which constitute the light take place in directions at right angles to the direction in which the light is travelling. This difference entails an entirely new set of phenomena, for whereas there is only one direction in which the wave is travelling, there are an infinite number at right angles to it. Consequently, the behaviour of light-waves must be associated with some direction in space at right angles to their line of propagation (i. e. in the wave-front), whereas, in longitudinal waves, there is nothing to distinguish any one direction in the wave-front from any other.

This fact may be readily shown in the case of the long artificial electromagnetic waves. If a grid of fine wires, wound parallel to one another on a wooden frame, be placed in front of a straight Hertzian vibrator, it is found that it stops the waves when the wires are parallel to the vibrator,

lets them freely through when the wires are at right angles to the vibrator. For convenience of description, let us suppose that the vibrator is placed horizontal in a N. and S. line, and that the grid is placed to the east of it, and on the same level: now the electromotive force generated in the vibrator is along its length, i. e. N. and S., and this is propagated through the surrounding space, keeping the same direction; consequently, when the waves fall on the grid, their direction of propagation is towards the east, and their direction of vibration is (so far as electromotive force is concerned) N. and S. If the vibrator be not moved, the direction of vibration remains the same for any length of time; when that is so the radiation is said to be *polarized*, and, to specify it precisely, a certain plane is chosen: this, by a convention based on quite other phenomena than those just described, is the plane at right angles to the electromotive force, i. e. the vertical plane. Accordingly, when radiation is stated to be polarized in a certain plane, this means that the electric oscillations it consists of take place at right angles to that plane.

Now ordinary light is not polarized, because the molecular vibrators from which it is derived are in all sorts of positions, and continually moving. It is, however, possible to sort out the waves by a device analogous to the metal grid, and any such device is called a polarizer: the simplest is a plate of tourmaline crystal. Such a plate allows light-waves to pass through it when their vibrations are in a certain direction relatively to the structure of the crystal, but, if they are at right angles to that direction, absorbs them; hence, whatever be the nature of the light before reaching the tourmaline, it is plane polarized afterwards. Then, if a second plate of tourmaline be put after the first, it will treat this polarized light just as the metal grid treats the radiation from the Hertzian vibrator; that is, if the second crystal plate is placed similarly to the first (so that the definite direction in the crystal, known as the optic axis, is the same in the two), it lets through the light coming from the first; but, if the axis of the second be put at right angles to that of the first, it is in a position to stop just that light which gets through the first, so that no light at all gets through the combination. The second plate is called an *analyser*, and, in the latter case, is said to be crossed to the polarizer. Various other devices for polarizing have been invented, and it may be mentioned here that it has even been found possible to construct a grid of such exceedingly fine wires that natural radiation of the longest wavelength (infra red) from hot bodies could be polarized with it.

### § 3. Phenomena of emission and absorption.

Light is usually produced by raising a solid to a high temperature; but, in order to examine precisely the transformations of energy accompanying this process, it is necessary to have



some instrument whose sensitiveness to radiation is not limited, as that of the eye is for physiological reasons, but is capable of appreciating indifferently radiation of all wave-lengths. Such an instrument is the thermometer, provided its working part be covered with a substance like lampblack, which absorbs all the radiation that falls on it, for then all the radiation is converted into heat, and the amount of it is registered by the rise in temperature. Thermometers far more delicate than the ordinary mercury in glass one have been invented for the purpose, into the details of which it is not necessary to go here. But again, the observations to be made on radiation would be very incomplete if it were not possible to examine separately that of different wave-lengths; to accomplish this a *spectroscope* must be used. Practical details of this instrument are given below (p. 294); here it may suffice to say that a narrow vertical slit is illuminated strongly by the light to be examined, and the slit viewed through a short telescope; but before reaching the telescope, the light is caused to pass through a triangular glass prism with its edge vertical. Such a prism bends the rays of light falling on it, and as the amount of bending depends on the refractive index of the glass, it is different for the various kinds of light, and these are dispersed. Thus, instead of seeing in the telescope a sharp white image of the slit, one sees a long coloured band, of the same vertical height as the slit, but changing laterally from red through the various colours of the spectrum to violet, the red rays being less bent than the blue. If an appropriate kind of thermometer be substituted for the eye-piece of the telescope, it can be moved about in the spectrum, and the latter can be traced into regions beyond the red. It should also be possible to trace it into regions beyond the violet at the other end of the visible band, but the radiation there is so faint that no thermometer sufficiently delicate has yet been constructed. Fortunately photography is of assistance here, so that our knowledge of the ultra-violet spectrum has mostly been obtained by putting a photographic plate in the place otherwise occupied by the eye-piece in a spectroscope.

Armed with these means of investigation, then, it is possible to determine the way in which radiation from a hot body depends on the nature and temperature of the latter.



Now it is a matter of common observation that a hot body, placed inside a cooler enclosure, eventually becomes of the same temperature as that enclosure. The equalization of temperature is accomplished partly or entirely by radiation: so long as the inside body is hotter than the enclosure it radiates heat to it; when equality of temperature is reached, the process apparently stops. That, however, is not a complete account of the phenomenon, for to suppose the radiation to stop would imply that whether a body radiates or not is regulated by the condition of some other body elsewhere—viz. the enclosure. In the words of Stokes ‘the molecules of the body cannot prophesy what is ultimately to become of the motion they may communicate to the ether, and regulate the communication accordingly.’ The true explanation is, therefore, that the body and the envelope both radiate to each other, under all circumstances, but when their temperatures are equal the amount of radiation is the same on both sides, and so no change of temperature results, but if the temperatures are unequal, the hotter body radiates more than it receives from the other, and consequently falls in temperature. This view is known as the *theory of exchanges*.

Experiments on the amount of radiation from hot bodies are best satisfied by Stefan’s formula (p. 75), according to which the radiation is proportional to the fourth power of the absolute temperature. This means that if two surfaces opposite one another are at different temperatures, say  $T_1, T_2$  (absolute), the energy radiated by the first to the second is equal to, say,  $kT_1^4$ , that radiated by the second to the first  $kT_2^4$ , and consequently the heat transferred by radiation from the hot to the cold body is the difference  $k(T_1^4 - T_2^4)$ . When, however, the difference of temperature is only a few degrees (e. g. between the human body and the walls of houses), Stefan’s law leads to the same result as Newton’s law of cooling (p. 74), that the loss of heat from a body by radiation is proportional to its excess of temperature above the surroundings.

The constant  $k$  for a perfectly black surface (e. g. lampblack) is about  $1.28 \times 10^{-12}$  calories per sq. cm. per sec. If this number be used to calculate the radiation from the outer to the inner surface of a vacuum jacket such as is used to contain liquid air, it will be found that the heat received by the liquid air is so

small that there is nothing anomalous in the fact that it evaporates very slowly. Actually, if the surfaces are made of glass, the radiation is a good deal less than the amount just mentioned.

To revert to the theory of exchanges, suppose a polished metal ball enclosed in a chamber with lampblack walls ; then we know that eventually the temperature of the ball will be the same as that of the enclosure. But now the radiation from the enclosure falling on the ball would balance that proceeding from the ball, if the latter were lampblack ; as it is the same amount falls on the ball, but some of it is reflected, and consequently only the remainder—say the fraction  $x$  of the whole—is absorbed ; yet this suffices to balance the radiation from the polished ball. It follows, then, that the radiation from the polished ball can only be  $x$  of that from the same ball lampblack, and at the same temperature. This result as to the relation between the powers of radiation and absorption is due to Kirchhoff. We may state it formally thus :—*If a surface absorbs the fraction  $x$  of the radiation falling on it, then it will radiate  $x$  times as much as a black surface at the same temperature.*

Accordingly, the constant  $k$  given above for lampblack must not be taken as correct for surfaces that reflect any of the rays that fall on it. For polished silver the radiation (at low temperatures) has been found to be only about 30 % that of lampblack, so that  $k = 0.3 \times 1.28 \times 10^{-12} = 3.84 \times 10^{-13}$  cal. per sq. cm. per sec.

Whilst, however, the total radiation from a solid when heated increases in proportion to the fourth power of the absolute temperature, it is not true that all the kinds of light increase in the same proportion. On the contrary, as already remarked, a solid at first gives out only radiation of very long wave-length, far too long to see ; as the temperature rises the intensity of this infra-red radiation is increased, but other kinds higher up the scale are added, until at a sufficiently high temperature the solid becomes luminous, i. e. radiation of wave-length shorter than  $7.6 \times 10^{-5}$  cm. is produced, so that the eye is affected by it. When the temperature is still further raised, the phenomena proceed in a similar manner : the intensity of the red as well as of the infra red is increased, but other colours, yellow, green, blue, &c., are added ; and always the increase in intensity is greater in the higher regions of the spectrum than



in the lower. The distribution of energy in the spectrum will be more easily grasped from the curves in Fig. 97. The height of the curve there represents the intensity of the radiation of the particular wave-length, as shown by the abscissae, and as

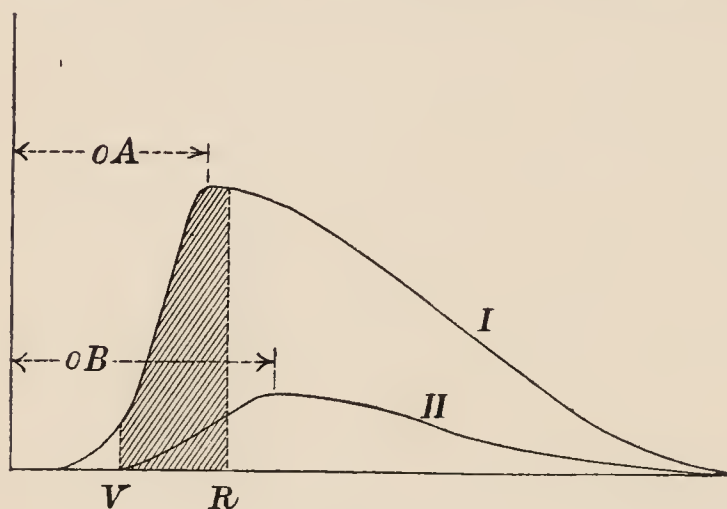


FIG. 97.

measured by putting a sensitive thermometer in the corresponding part of the actual spectrum. Curve I represents approximately, according to Langley's measurements, the solar radiation; curve II, that from a source at low temperature; and it will be noticed that

not only is I much more intense than II in all parts, but that the difference becomes greater as we pass from the long wave-lengths to short, and that consequently the wave-length ( $oA$ ) for which the maximum of radiation occurs is shorter in sunlight than that ( $oB$ ) for the other source. Further, as a consequence of this, the fraction of the total radiation which occurs between the limits ( $VR$ ) of the visible spectrum (represented by shading), is much greater for sunlight than for the low temperature source. Even for sunlight it is only about 36 per cent., so that regarding the sun as an illuminating appliance, more than half its energy is useless; while for such a source as an incandescent lamp only 2 or 3 per cent. of the total radiation lies within the limits violet to red, and all the rest is wasted.

What precedes is an account of the normal production of radiation as a consequence of temperature—what may conveniently be known by the name of *incandescence*. But it very often happens that substances, for one cause or another, give out radiation of a kind very different from that which their temperatures would indicate, becoming luminous, for example, at temperatures far below a red-heat; these cases are included under the general term *luminescence*. The most important causes which give rise to this phenomena are (i) chemical, (ii) electrical, (iii) action of light. Chemi-luminescence occurs in many instances, of which the phosphor-



escence of phosphorus may be taken as example. It is probable that this is due to a slow oxidation, but it is at any rate certain that the phosphorus, while giving a whitish glimmer, remains cold. Probably the luminous phenomena of the glow-worm, and other animal organisms, come under the same head. Electro-luminescence is very familiar in the discharge of electricity through gases: vacuum tubes can be made to glow with considerable intensity by a discharge, but the temperature reached by their contents is by no means high. Photo-luminescence may be taken as including all the cases in which substances exposed to light become phosphorescent, i. e. capable of giving off light on their own account afterwards (e. g. calcium sulphide, or Balmain's luminous paint); phosphorescence is really a very common phenomenon, but in a much slighter degree, that requires special appliances to detect it. Finally, the important case of *fluorescence* comes under the same head: there are substances (e. g. fluorescein) which send out light on their own account, so long as they are illuminated by an independent source, but not afterwards; thus, if a solution of sulphate of quinine be illuminated by violet light, it will be seen, looked at sideways to the light, to be of a pale blue colour, as if self-luminous. A solution of chlorophyll in alcohol, similarly treated with green light, appears red. The fluorescent substance absorbs some particular kind of light falling on it—violet in the one case, green in the other, and, in consequence, becomes luminous, giving out in all directions light, not of the same kind, but always lower down in the spectrum—blue and red respectively in the instances quoted. It appears to be in somewhat the same way that Röntgen rays, falling on a fluorescent screen, make it self-luminous, so that if an organ be placed between the source of the rays and the screen, those parts of the organ which most obstruct the rays, such as bones, appear in shadow on the screen.

Solids and liquids, when made incandescent, all give a continuous spectrum, i. e. one including all possible kinds of light within certain limits of wave-length; so that their spectra are not characteristic of the substances producing them, but merely of the temperatures. Gases, on the other hand, when not under too high pressure, give out spectra consisting of light of a few definite kinds only, and therefore appearing in the spectroscope as a series of bright lines (images of the slit) in definite positions, with dark spaces between. The explanation of this difference on the molecular theory, is that the molecules of a gas, being widely spaced, are able to travel quite appreciable distances without coming into collision, or being affected by neighbouring molecules; hence, after a collision has set a molecule in vigorous vibration, it may execute a thousand, or even

a million, oscillations before it is again interfered with by another collision. During this interval it gives out waves of its own natural period to the ether, just as a tuning-fork struck, and then after five minutes struck again, would give out its natural sound to the air, the time occupied in striking it being negligibly small compared to the intervals in which it is left alone. Whereas in solids and liquids a molecule never escapes from the influence of its neighbours, so that it has no opportunity of vibrating in its own natural period. The transition between the two may be seen by observing the spectrum of a gas (produced by an electric discharge through it), and gradually increasing the density of the gas. The 'lines' in the spectrum widen out into bands, these spread more and more, till they fill up all the intervening dark spaces, and a continuous spectrum is formed.

Since the spectrum of each gas consists only of a few bright lines (or in some cases bands), occupying fixed positions in the scale of wave-lengths, and answering to free periods of vibration of the gas molecules, that of each gas is different to all others, and the spectroscope affords a means—highly delicate and certain—of detecting particular gases. Thus, hydrogen gives the three bright lines referred to on p. 261, the position of which can be measured with an accuracy approaching one part in a million; sodium vapour gives two lines in the yellow, called  $D_1$  and  $D_2$ , of wave-length  $5.896 \times 10^{-5}$  and  $5.890 \times 10^{-5}$  cms. respectively. These two appear as one in a spectroscope of small dispersion, and are known as the D line; and since they are very bright and easy to get, by merely placing a sodium salt in a bunsen burner, refractive indices and other optical constants are very commonly measured with the aid of it. Thus, the standard refractive indices given in the table, p. 264, and called  $\mu_D$ , are the values of  $\mu$  for this particular kind of light.

In order to obtain a gaseous spectrum for examination, the substance to be studied may be volatilized in a bunsen flame, or if that is not hot enough, in an arc lamp; or, if a metal, electrodes may be made of it, and a spark from an induction coil passed between them, when it will volatilize some of the metal, and take a characteristic colour accordingly; or a vacuum tube may be prepared, containing a little of the substance, and a discharge passed through it, rendering it luminous. The latter method



has the advantage that, the gas being rarefied, the molecules are more completely free to execute undisturbed vibrations than when at atmospheric pressure, and the lines of the spectrum are therefore sharper.

By examining the light from nebular comets, and the outer fringe of the sun as it appears in total eclipses, it has been shown that these heavenly bodies are gaseous, and the actual gases they consist of have been determined.

Kirchoff's law as to the equality between emission and absorption is true not merely for the radiation as a whole, but for each kind separately; consequently, substances which possess special powers of emitting particular rays of light, possess special powers of absorbing the same rays. Hence arise what are known as absorption spectra, which may be as characteristic of a substance as the emission spectra. If white light from a very hot source, such as an arc lamp, be allowed to pass through a bunsen burner made yellow with sodium, and the light be then examined with a spectroscope, it is found that the continuous spectrum of the arc is interrupted by two dark lines, precisely in the positions in which the bright yellow lines would be seen, if the sodium burner were looked at alone. These lines are not really dark, but appear so because they are less luminous than the neighbouring parts of the spectrum. They arise in this way: sodium vapour is transparent to most kinds of light, so that the arc spectrum as a whole is unchanged by the flame; but when waves of the particular frequency of the D lines fall upon sodium molecules, they set those molecules vibrating in precisely the same way that an acoustical resonator is set vibrating, when waves in unison with its own natural note pass over it. All that has been said (p. 157) about resonance in sound is applicable here; the energy of the incoming light-waves of the right frequency is absorbed, i. e. it is spent in setting the molecules of the absorbing substance (sodium vapour) in vibration, and consequently does not pass on and reach the spectroscope. However, the sodium molecules themselves in this way become a source of waves (in addition to the effect produced by the heat of the gas flame), and give out light, not only, of course, in the direction of the spectroscope, but in all directions. Some light of the given frequency therefore reaches the spectroscope, and it is a mere question of temperature whether the D lines



appear dark on a bright spectrum, or brighter than the rest of the spectrum. With an arc as source of light, the radiation is so intense that the flame absorbs more than it gives out on its own account; but with an incandescent lamp as source, the radiation is weaker, and the flame supplies more yellow light than it absorbs, so that the lines appear bright.

The spectra of the sun and stars show a continuous band of light interrupted by fine lines (the Fraunhofer lines), which are now known to be due to the white light from the core of the sun or star being partly absorbed by gases in the atmosphere above it—partly too by the earth's atmosphere. Very many of the lines have been identified in position with bright lines in the emission spectra of known bodies, and in this way the existence of many elements known on earth has been proved in the sun and stars. Even an element called helium was known for many years in the sun, being characterized by a line that did not agree with those of any known substance, and, subsequently, that element was discovered by Ramsay in certain terrestrial minerals.

A few of the most prominent of the Fraunhofer lines have received letters as a designation, and they form a most convenient way of describing the different parts of the spectrum. Their position is sufficiently indicated by the accompanying chart (Fig. 97 a): as already mentioned, C and F are due to hydrogen, D to sodium.

Solids and liquids, though never emitting light of sharply defined character like gases, sometimes emit light belonging to one region of the spectrum more strongly than another; the corresponding fact, absorption of one part and not another, is familiar—it is in fact the

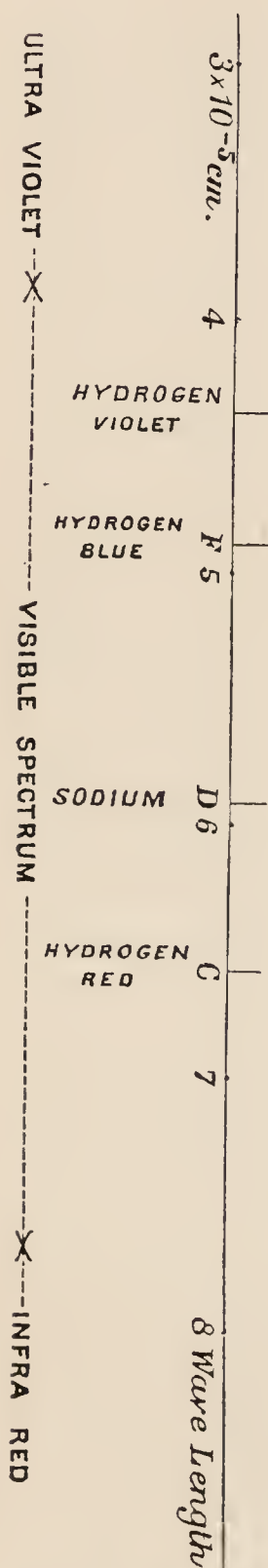


FIG. 97 a.

cause of the colour of ordinary objects. Thus, if a cell containing bichromate solution be placed in front of the slit of a spectro-scope, it will be found that the spectrum of a source of white light is reduced to the red end only; the solution is transparent to red light, but opaque to (i. e. absorbs) other colours. As a rule, in solids and liquids, such selective absorption extends in a somewhat indefinite manner to large tracts of the spectrum, but in a few cases it consists in bands of fairly definite position. Thus, haemoglobin produces two dark bands lying between D and E in the yellow and green. Hence, the appearance of these bands may be used as a delicate test for the presence of blood.

It should be noted that selective absorption may occur just as well in the ultra violet or infra red: a body in which this occurs appears colourless, but in the physical sense it is similar to a coloured body, e. g. glass absorbs the ultra violet strongly; hence, if we obtained visual sensations from the ultra violet, it would not appear perfectly transparent and colourless. Quartz, on the other hand, transmits ultra-violet light freely. Water is opaque for radiation of wave-length of more than about  $14 \times 10^{-5}$  cm. The limitation of the eye to radiation between  $7.5$  and  $3.8 \times 10^{-5}$  cm. may be due either to a physical or a physiological cause: the radiation may be absorbed before it reaches the retina, or it may be without effect on the retina when it does reach it. The former appears to be the cause of the limit at the violet end, but is certainly not so at the red end, as it has been shown that longer waves can traverse the media of the eye.

#### § 4. Reflection and refraction.

Very many optical phenomena, and especially those involved in the action of mirrors, lenses, and other optical instruments, can be determined with sufficient approximation by the assumption that light consists of rays which, when travelling in a uniform medium, are straight, and which, at the boundary of two media, suffer reflection and refraction according to the laws stated on pp. 266, 267. This method of treatment yields what is known as *geometrical* optics; being based only on rules which give an approximation to the facts, it is occasionally insufficient, e. g. the theory of the resolving power of the microscope, due to Abbe,

though of great practical importance in the construction and use of that instrument, lies outside the domain of geometrical optics. But we shall not attempt any such refinements here, it being sufficient for most purposes to treat optical instruments, including the eye, by means of the rules of reflection and refraction.

The application of these rules consists largely in determining the position of 'images' formed by mirrors and refracting surfaces. The precise meaning of this term will be best understood after considering the simplest instance in detail.

*Reflection at a plane surface.* Let  $A$  (Fig. 98) be a point source of light,  $BCD$  a plane reflecting surface; draw the perpendicular  $AB$  and produce it to  $A'$ . Then a ray of light from  $A$ , falling on the mirror at  $B$ , since it is normal to the surface, will be reflected

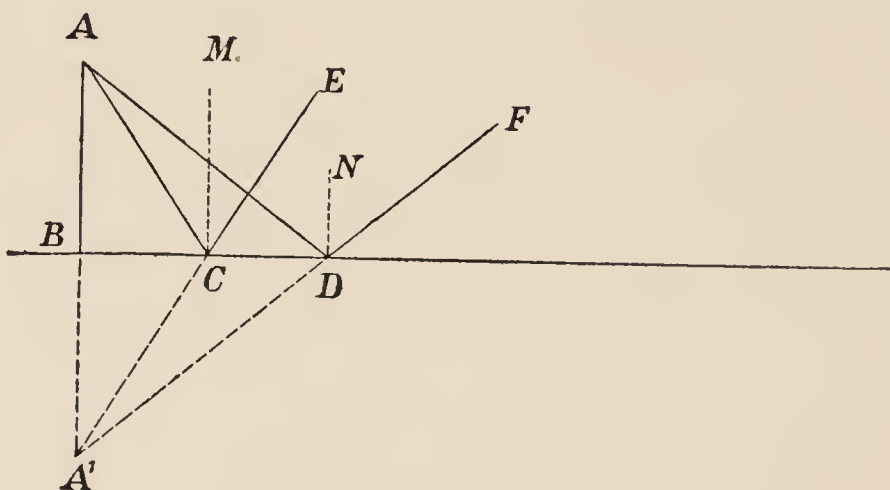


FIG. 98.

back along the same line. Draw another ray  $AC$ , and at  $c$  erect the normal  $CM$ ; draw  $CE$  on the opposite side of the normal, so that  $CA$  and  $CE$  make the same angle with the normal; therefore  $CE$  is the course of the ray reflected at  $c$ . Produce  $CE$  backwards to meet the production of  $AB$  at  $A'$ ; then  $A'B$  is equal to  $AB$ . For the triangles  $ABC$ ,  $A'BC$  have equal angles and a common side; therefore they are equal in all respects. Now the lines of the two reflected rays  $BA$  and  $CE$  intersect in  $A'$ ; and, to an eye placed in the direction of  $A$  and  $E$ , they will appear to come from  $A'$ . That, however, will not be true of the whole beam, unless all the rays, after reflection, are in lines passing through  $A'$ . Consider, then, some other ray,  $AD$ , and in a similar way draw the normal  $DN$  and the reflected ray  $DF$ ; then, by a similar argument, it follows that  $DF$ , produced back-



wards, passes through  $A'$ , and so on for all the rays. Hence, to an eye placed, say, in the direction of  $E$  and  $F$ , all the light after reflection appears to come from  $A'$ , and it is immaterial whether it comes from a source at  $A'$  through a window  $BCD$ , or from a source at  $A$  by means of a mirror  $BCD$ .  $A'$  is called the *image* of  $A$ , and we may give the following formal definition—

When, after reflection or refraction, all the light proceeding from one point proceeds to, or appears to proceed from, a second point, the latter is called an image of the former; if the light actually passes through the second point it is called a *real* image, otherwise a *virtual* one.

A plane mirror, therefore, forms a virtual image. This fact may easily be verified by holding a candle in front of a looking-glass.

*Refraction at a plane surface.* Let  $A$  (Fig. 99) be a point-source of light,  $BCD$  the bounding surface between air and a second medium. The course of any ray  $AC$  may be found by drawing the normal  $MCM'$ , then the angle  $ECM'$ , such that

$$\sin ACM = \mu \sin ECM',$$

where  $\mu$  is the refractive index of the second medium; then  $CE$  is the direction of the refracted ray. It will be found that the refracted rays  $CE$ ,  $DF$ , &c., diverge very approximately from the point  $A'$ , which is on the normal  $AB$ , and at a distance from the surface at  $B$  such that  $AB = \mu A'B$ .

The image formed in this case is only approximate, whereas in the former case the lines of all the rays passed exactly through  $A'$ . Departure from exactness in the formation of images is called *aberration*, and occurs in nearly all optical instruments to a greater or smaller extent.

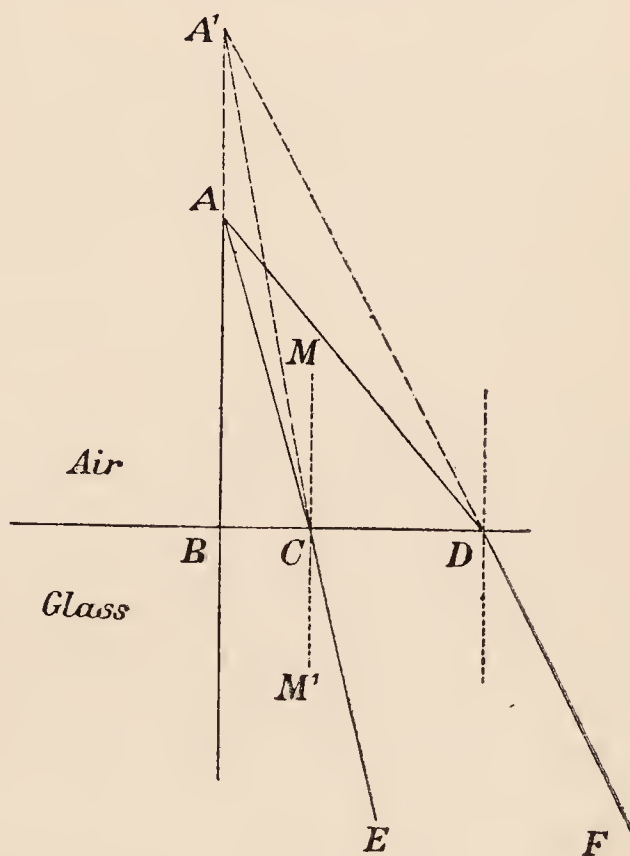


FIG. 99.

*Reflection at a spherical surface.* In this and the following cases also the formation of an image is only approximate, but is sufficiently close for most purposes, when the surface does not form too large an arc of a sphere, e. g. an ophthalmoscope mirror consists of a section of a sphere not more than a few degrees across; if it were a hemisphere, or even a quarter sphere, the aberration would cause great distortion in the appearance of the objects looked at. Two problems arise, (i) to find the position of the image of a point, (ii) to find the size of the image of an object of finite size—this being made up of the images of all the points that compose the object. With regard to the former, it will be sufficient to consider points on the axis, i. e. the normal to the mirror through its middle point. A geometrical construction for this is given in Fig. 100 for the case of a concave mirror. Let

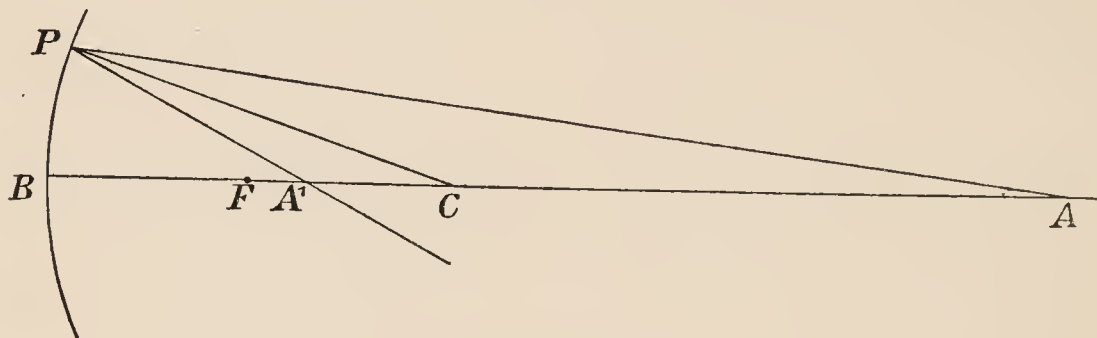


FIG. 100.

A be the source of light, on the axis ACB; since the axis is a normal to the mirror, it must pass through the centre of the sphere, let this be C. Draw a ray AP and the normal CP, and make the angle  $CPA' = CPA$ ; then  $PA'$  is the course of the ray reflected at P. It may be shown that all the rays after reflection cut the axis at the same point  $A'$ , which is consequently the image, and in this case a real one. The most convenient way of finding the position of  $A'$  is, however, an arithmetical one that will be explained below.

The position of the image varies with that of the object, certain cases being especially important. If the object is at the centre of the sphere the image coincides with it. If the object is very far away (mathematically speaking at infinity) the image is at F halfway between C and the mirror at B—this position is known as the *principal focus*; and if the object is at the principal focus the image is at infinity. Further, if the object is nearer

to the mirror than  $F$ , the image is virtual and on the opposite side of the mirror. The distance  $BF$  is called the *focal length*, and, in the case of a mirror—concave or convex (but not of a refracting surface)—is clearly half the radius of curvature. All these points should be verified by drawing the corresponding diagrams, and the convex mirror should be treated similarly.

Assuming then the proof, for which there is not space here, that a spherical mirror does bring to a point-image all the rays falling on it from a point-source, it is very easy to find the position of the image of a point off the axis. For to do so we have only to trace the course of two rays: where they intersect we know that the other rays must intersect too. The geometrical method is given in Fig. 101. From  $H$  draw two rays, the first,  $HP$ , through the centre  $C$  of the mirror; then, since it falls

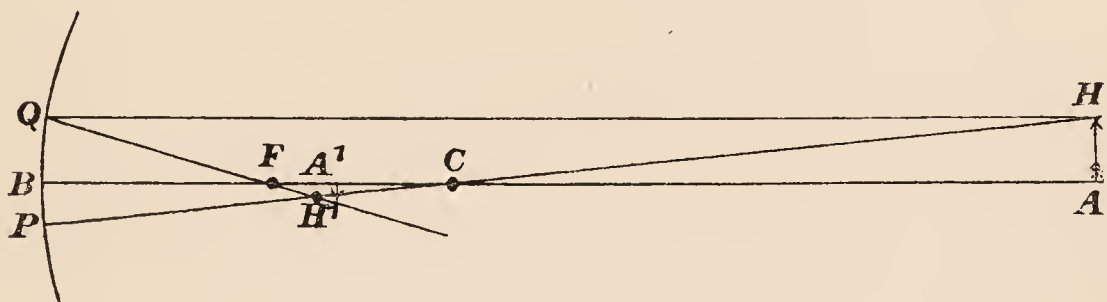


FIG. 101.

normally on the mirror, it must be reflected along the same line; the other,  $HQ$ , parallel to the axis  $ACB$ ; then, since this might be part of the rays coming from an object indefinitely far away, it must be reflected through the principal focus  $F$ , so that its direction after reflection can be found with a pair of dividers and a ruler. The intersection  $H'$  is the image of  $H$ .

Further, by dropping perpendiculars from  $H$  and  $H'$  on to the axis at  $A$  and  $A'$ , the size of the object and image may be compared; for if there be an object of finite size  $AH$ , then  $A'H'$  will be the image of it, and we see that *when object and image are on the same side of the centre, the image is upright*, when (as in the figure), they are on opposite sides, the image is inverted; and that (seeing the triangles  $AHC$  and  $A'H'C$  are similar)  $A'H' : AH :: A'C : AC$ , or

$$\frac{\text{size of image}}{\text{size of object}} = \frac{\text{distance of image from centre}}{\text{distance of object from centre}}.$$



This ratio is called the *magnification* produced by the mirror. It may be (as in the figure) less than unity, but is still called magnification.

The student should draw corresponding figures with the object in other positions, and also for a convex mirror.

*Refraction at a single spherical surface.* A lens has two refracting surfaces, but the case of a single surface is not only important as leading up to the other, but it gives a good first approximation to the action of the eye. The 'reduced eye' of Listing, which reproduces fairly the really complex optical structure of the eye, consists of a single surface, between air and a medium such as the aqueous humour, convex towards the air, and of radius 5.1248 mm. We shall take this as example. In Fig. 102 BP is the curved surface, AB the axis, A a luminous point on the axis, C the centre of curvature, CPN consequently the normal

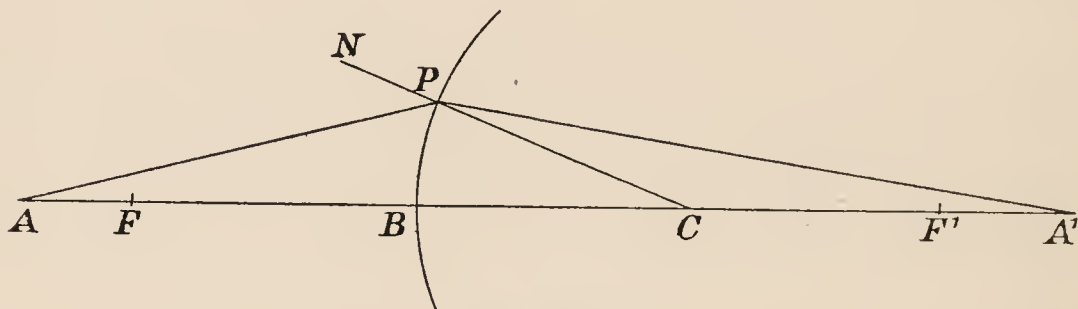


FIG. 102.

at P. An incident ray AP makes the angle of incidence APN in air; the angle of refraction in the aqueous humour CPA' is such that  $\sin APN = \mu \sin CPA'$ , where  $\mu$ , the refractive index, is 1.3376. The refracted ray is PA' cutting the axis at A', so that this point is the image (real) of A. If light comes from infinitely far away on the axis, on the air side, it is brought to an image at F', which, since it corresponds to an infinitely distant source, is called the principal focus. Again, if light started from a point indefinitely far away in the aqueous humour, it would form an image at F; this is, therefore, also a principal focus, and to distinguish the two, F is called anterior, F' posterior. The positions of F and F' can be found geometrically, but we shall subsequently give a convenient method of finding them arithmetically.

Now suppose that the source of light is a small object AH

(Fig. 103), so that it includes points off the axis. To find the image of  $H$  two rays must be traced out, and the same rays as before are the most convenient; one  $HCH'$  passing through the centre, since it falls normally on the refracting surface, is undeviated; another,  $HP$ , parallel to the axis is refracted through

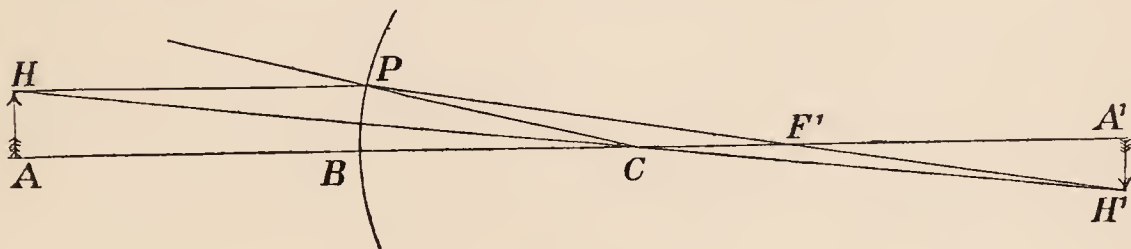


FIG. 103.

the posterior principal focus  $F'$ , so that its course is  $PF'H'$ ; the intersection  $H'$  is the image of  $H$ . Consequently  $A'H'$  is the image of the object  $AH$ , and the same rules hold for magnification as in the foregoing case.

*Numerical treatment of lenses and mirrors.* It remains to show how to deal in a convenient numerical manner with the instances of reflection and refraction described above. For this certain definitions must be introduced. We have already referred to the radius of curvature of a surface, i. e. the radius of the sphere of which the surface forms part; but it is clear that the larger the radius is, the flatter the surface will be. Accordingly, it is more useful to state the *curvature*, defining this as the reciprocal of the radius of curvature. To measure this a unit is required, which is the reciprocal of a length: the unit in use is the reciprocal of a metre, and is called a *dioptr* (written  $D$ ), e. g. the surface of the reduced eye has a radius of 0.0051248 metres, so that its curvature is  $1 \div 0.0051248 = 195.1 D$ .

Again, a beam of light is in general convergent or divergent, and clearly—to take the former case—it is more strongly convergent when converging towards a point near at hand than one far away. It is best to take as a measure of the *convergence* the refractive index of the medium in which the light is travelling  $\div$  distance of the point to which the light is converging, and to treat divergence as negative convergence; these quantities can also, then, be measured in dioptr. The reason for this definition lies in its convenience, as will best be understood from what follows. By way of example, in Fig. 102 the light is supposed to



start from the point A at a distance of 0.1 metre from the refracting surface; consequently, when it reaches that surface, it possesses a divergence of  $1 \div 0.1 = 10$  diopters, or a convergence of  $-10$  D.

Now it can be shown that a spherical surface, whatever be the divergence or convergence of the light falling on it, alters that divergence or convergence by a fixed amount. We therefore give the following definition:—*The strength of a spherical reflecting or refracting surface is the amount by which it increases the convergence of light falling upon it.* By the aid of the quantities here defined it is extremely easy to trace out the changes produced in a beam of light by any number of surfaces. The actual strengths of the various surfaces are as follows:—

Plane reflecting or refracting surface . . . . .	0
Concave mirror . . . . .	$2\rho$
Convex mirror . . . . .	$-2\rho$
Spherical refracting surface convex towards the air .	$(\mu - 1)\rho$
Do. concave towards the air . . . . .	$(1 - \mu)\rho$

Here  $\rho$  is the curvature of the surface,  $\mu$  the refractive index of the medium on one side of the refracting surface, it being supposed that there is air on the other side<sup>1</sup>.

The strength of the reduced eye is consequently  $0.3376 \times 195.1 = 65.8$  diopters. In Fig. 102 the beam falling on the eye has a convergence, as shown above, of  $-10$  D; add to this the strength of the surface and we arrive at  $+55.8$  D as the convergence of the refracted beam on leaving the surface. Hence, by the definition of convergence, it is converging towards a point (image) at a distance  $\mu \div 55.8 = 0.02398$  metres from the surface: this is the point A'.

It is still simpler to find the principal focal points F and F'. If a beam of light come from a very distant object on the air side, its convergence on reaching the refracting surface is 0; consequently, on leaving that surface it is 65.8 D, and it will form a focus at  $1.3376 \div 65.8 = 0.02030$  metres or 20.30 mm.

<sup>1</sup> When there are two media other than air separated by the surface, the strength is  $(\mu_1 - \mu_2)\rho$ ,  $\mu_1$  being the refractive index on the concave side,  $\mu_2$  on the convex side of the surface. This formula is not used in the examples given in the text, but is of great practical importance, e. g. in the case of a water immersion objective.



from B. The *posterior focal length*  $BF'$  is, therefore, 20.30 mm. The retina is situated at  $F'$  so that objects at a great distance off form images on the retina. The object in Fig. 103 would form, as we have seen, its image 23.98 mm. from B and consequently 3.68 mm. behind the retina. The actual eye possesses a power of accommodation to bring the images of near objects on to the retina.

If a beam of light could start from indefinitely far away on the side of the aqueous humour in the reduced eye, it would fall on the refracting surface with zero convergence and come out into the air with convergence 65.8. But as it is now travelling in a medium of refractive index unity, the point to which it converges is at a distance  $1 \div 65.8 = 0.01518$  metres or 15.18 mm. away. This is the *anterior focal length*  $BF$ .

## § 5. Optical instruments.

*Thin lens.* A lens is a combination of two refracting surfaces, either of which may be convex, plane, or concave. Usually the two surfaces are but a short distance apart; this alone, however, does not constitute a thin lens in the technical sense of the word: it is necessary that the distance between the two faces should be small compared with their radii. This condition is commonly, but not always, satisfied by the lenses in actual use; the higher power objectives of a microscope contain lenses which, though very small actually, approach a hemisphere in shape, and consequently cannot be treated as thin.

The action of a thin lens can be represented by its total strength, which is the algebraic sum of the strengths of the two refracting faces. Thus in a bi-convex lens, i. e. one in which both surfaces curve outwards, if  $\rho$  and  $\sigma$  are the curvatures, the strength is  $(\mu - 1)(\rho + \sigma)$ ; in a lens convex one side, with curvature  $\rho$ , concave the other, with curvature  $\sigma$ , the strength would be  $(\mu - 1)(\rho - \sigma)$ .

A simple rule, based on the effect of curvature, indicates whether the strength of a lens is on the whole positive or negative: a thin lens is convergent (+<sup>ve</sup>) when it is thickest in the middle, divergent (-<sup>ve</sup>) when thinnest in the middle.

We may describe the algebraic sum of the curvatures as the

total curvature  $\tau$  and write the strength as  $(\mu - 1)\tau$ ; the quantity  $\mu - 1$  is sometimes called the refractivity, hence in words—

strength of a lens = refractivity  $\times$  total curvature.

A lens has an anterior and posterior principal focus, but since there is the same medium, air, on both sides of it, these are at equal distances from it, viz.  $1 \div \text{strength}$ . This is known simply as the focal length of the lens. It can easily be found in the case of a convergent lens by holding it up to sunlight; a very bright spot will be formed at the principal focus on the side away from the sun, which can be detected by a sheet of paper. In accordance with the recommendation of an ophthalmological congress, spectacle lenses are now classed by their strengths; thus a lens of  $+2$  D is a convergent one of  $\frac{1}{2}$  metre focal length; a lens of  $-4$  D is a divergent one of  $\frac{1}{4}$  metre focal length.

A thin lens possesses a point called the *optical centre*, such that

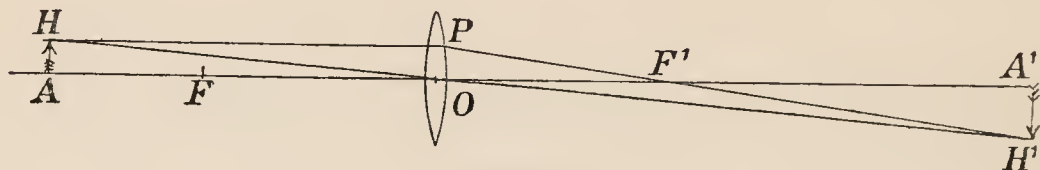


FIG. 104.

rays passing through it are undeviated; this is not the centre of curvature of either face, but practically a point on the axis of the lens about halfway between the two faces. This point plays the same part in the action of the lens, as the centre does in a single reflecting or refracting surface. Hence, to find the image of a point H off the axis (Fig. 104), we trace two rays, one HOH' through the optical centre O, the other HP parallel to the axis, which on refraction passes through the principal focus F', becoming PF'H'; the point of intersection H' is the image. Consequently a finite object AH gives an image A'H'; the rule of magnification is the same as before, distances being reckoned from the optical centre.

The diagram should be drawn for the case of A being nearer to the lens than F, the anterior principal focus, when the image will be found to be virtual; and also for a divergent lens, which always gives a virtual image.

The chief uses of a single lens are (i) to form a real image, as in the objective of a microscope or telescope; (ii) to form

a magnified virtual image of a small object, as in the hand magnifier; (iii) to bring the divergence of a beam of light within the range of accommodation of the eye, as in spectacle lenses.

The first case corresponds to Fig. 104. If the object is infinitely far away (astronomical telescope), the image is at the posterior principal focus. If the object is nearer (laboratory telescope), the image is formed further away. In the microscope an objective of very short focal length is used, and the object brought so close to it, that the image is many times further away, and consequently many times larger; still the object cannot be brought so close as the anterior principal focus, else the image would be formed infinitely far away. E. g. an objective of 100 D is required to form a real image at a distance of 16 cms., that being the length of the tube. The convergence of the light, on leaving the lens, in order to meet at a point 0.16 metre away, must be  $1 \div 0.16 = 6$  D nearly. Hence, before reaching the lens, its convergence must have been  $-94$  D, and the light is diverging and started from a point whose distance is  $1 \div 94 = 0.0106$  metre or 10.6 mm., the focal length being  $1 \div 100 = 0.01$  metre or 10 mm.

In the second case, it is usually desired to form a virtual image of the small object at a distance from the eye that is

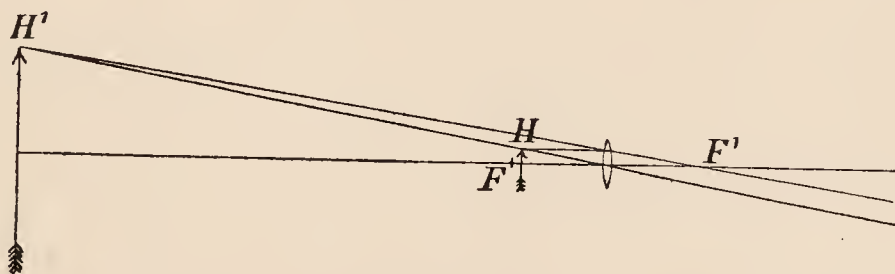


FIG. 105.

convenient for vision—say  $\frac{1}{5}$  metre. This means that the light when it reaches the eye must be divergent, its convergence being  $-5$  D. The lens is held close to the eye, consequently, if  $x$  is the convergence of the light before reaching the lens, and  $S$  the strength of the lens, we must have  $x + S$ , the convergence after passing the lens  $= -5$  or  $x = -(S + 5)$ . This means that it is divergent, and from a point at a distance  $1 \div (S + 5)$  metres away. Now the magnification or ratio of the size of the image to that of the object is the same as the ratio of their distances from



the lens, i. e.  $\frac{1}{5} : \frac{1}{S+5}$ , or  $S+5 : 5$ , or  $\frac{S}{5} + 1$ , e. g. a 60 D lens used as an eye-lens would magnify  $60 \div 5 + 1 = 13$  times. The course of the rays is shown in Fig. 105.

The third case will be considered in detail in connexion with the eye.

*Eye-pieces.* A simple magnifying lens suffers from two defects: (i) spherical aberration, already touched on briefly; the rays from a point in the object are not brought accurately to a point-image, and moreover, the image of an object of finite size suffers distortion; (ii) chromatic aberration, i. e. the lights of different colours (wave-length) have different refractive indices, so that the focal length of the lens is different for each colour; it is therefore at best only possible to adjust the lens so as to form an image with the average light of the spectrum, and the image will have coloured edges. A lens or system of lenses that did not suffer from the first defect would be called *aplanatic*, one that was free from the second, *achromatic*. It is not possible to arrive at these conditions precisely, but by a combination of two (or more) lenses, an eye-piece can be formed which is far more aplanatic and achromatic than a simple eye-lens. The two combinations that have met with the most success in practice are Huyghens' and Ramsden's.

Ramsden's (positive) eye-piece is mostly used for telescopes. It consists of two thin plane-convex lenses, of equal focal length (say 25 mm.), placed with their convex faces towards each other, at a distance apart equal to two-thirds of the focal length of either lens. The combination has a focal length of  $\frac{1}{4}$  that of either lens, so that if an object be placed at that distance (say  $6\frac{1}{4}$  mm.) from it, the rays from the object, after passing the two lenses, will emerge in a parallel beam, and can be dealt with by the eye. If the object be put slightly closer, the rays on emerging will be slightly divergent, appearing to come from a virtual image, as in Fig. 105, for the simple lens. A pair of cross wires, or a finely divided scale, may be set up in the position indicated, and will be seen by the observer. Of the two lenses that nearer the eye is called the eye-lens, the other the field-lens.

Huyghens' (negative) eye-piece is mostly used for microscopes. It consists of a small plano-convex eye-lens (say 13.3 mm. = 75 D),

and a larger plano-convex field-lens of three times the focal length (40 mm. = 25 D); these are fixed at a distance apart equal to two-thirds the focal length of the field-lens (26.7 mm.), and the flat side of each is turned towards the eye. It does not give a real image outside the eye-piece, and so cannot be used, like Ramsden's, to examine a small object directly—for this reason it is called negative. But if placed so that light converging towards a point 20 mm. behind the field-lens fall on the latter, it follows the course shown in Fig. 106, and emerges



FIG. 106.

in a parallel beam. In traversing the eye-piece the light forms a real image halfway between the two lenses, and then diverges again. Hence, if cross wires or a micrometer be used, they must be placed here.

*Telescope.* The astronomical or laboratory telescope consists of an objective combined with a Ramsden eye-piece. The objective is a large convex lens of considerable focal length, and in order to make it achromatic it is usually made in two parts, a bi-convex lens of crown glass, and a concavo-convex lens of flint, the latter being more concave than convex, so as to be divergent on the whole. The concave side fits against the crown lens, and the two may be cemented together with Canada balsam. The objective forms a real image, and the length of the telescope tube is so adjusted that the eye-piece can be set in front of this image, so as to view it like an actual object. For this purpose the tube of the telescope is made in three parts, and the mode of adjustment is: Slide the eye-piece tube in the middle tube, till the cross wires, which are carried by the latter, are clearly seen; then slide the middle tube (carrying the eye-piece with it) in the objective tube, till the object is seen without effort at the same time as the cross wires: when that is the case, the real image and the cross wires are in the same plane. The magnifying power of the instrument is measured by the ratio of the focal length of the objective to that of the eye-piece; but the resolving power, i. e. power of making out detail in the object looked at, depends rather on the width of the objective, as a wide lens collects more light than a small one.



*Microscopes.* The ordinary compound microscope consists of an objective and a Huyghens' eye-piece. The objective, which is very small in diameter but of great strength (i. e. short focus), is usually built up of several plano-convex lenses, with their plane faces away from the eye, to avoid spherical aberration; and each piece may be a pair of lenses of crown and flint, so as to be achromatic. The body of the microscope is moved by a screw, so as to come to the necessary distance from the object. This, as already stated, is such that the object lies slightly beyond the (anterior) principal focus of the objective. That lens would therefore form a real image, near the top of the tube, and greatly larger than the object; but the eye-piece is interposed before the rays meet, so that they may traverse it in the manner required to emerge in a parallel or slightly divergent beam convenient to the eye.

The magnifying power of a microscope depends chiefly on the shortness of focus (i. e. strength) of the objective; but, as in the telescope, it is no good to pay attention to this condition alone, for the resolving power, and hence the excellence of the pictures afforded by the microscope, depends on the width of the beam of light entering the objective. If in Fig. 107 A be

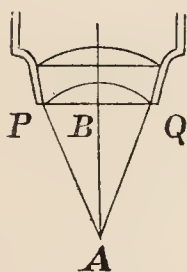


FIG. 107.

a minute object in view before the objective PBQ the angle PAQ includes all the light which enters the microscope and is effective. The extent of this beam is called the *aperture* of the lens, and the best way of reckoning it is by what is called the numerical aperture (N.A.), which is the sine of half this angle, that is to say  $BP \div AP$ . The numerical aperture obviously could never be greater than unity, and practical difficulties of construction limit it to about 0.9. This is when the object is in air; but if the space between the object and the objective is filled up with water, or preferably a liquid (cedar oil, monobromonaphthalene) of about the same refractive index as the glass of the lens, the resolving power is improved, and so the N.A. of a lens is defined practically as the ratio  $BP \div AP$  multiplied by the refractive index of the medium intended to be used with it, and may then be taken as a measure of the performance of the lens.

*Photographic camera.* This consists, optically, of an objective



merely provided with a screen on which to form a real image. A telescope or microscope objective can sometimes be used with success for the purpose, the latter for photographing minute objects, but two special features are required in a good photographic objective. (i) A telescope objective has only a very narrow field, but a photographic camera is expected to form a picture covering a wide area, which must all be accurately focussed together; the correction of the lenses for spherical aberration must therefore be such, that even points far away from the axis form point-images, and that the image of the whole should be flat and free from distortion. This is a much more difficult problem than that of the telescope, and to accomplish it four, six, or even eight lenses are sometimes used to build up the complete objective. (ii) The light that is brightest to the eye is the yellow-green, and therefore, in ordinary optical instruments, attention is paid especially to achromatization round this as a mean. But in a photographic camera the lenses should be adjusted more especially with regard to the blue-violet, since that is the light which is most effective in producing the chemical action on which photography depends.

### § 6. Dispersion.

The decomposition of white light into the set of colours known as the spectrum is usually accomplished by a triangular prism

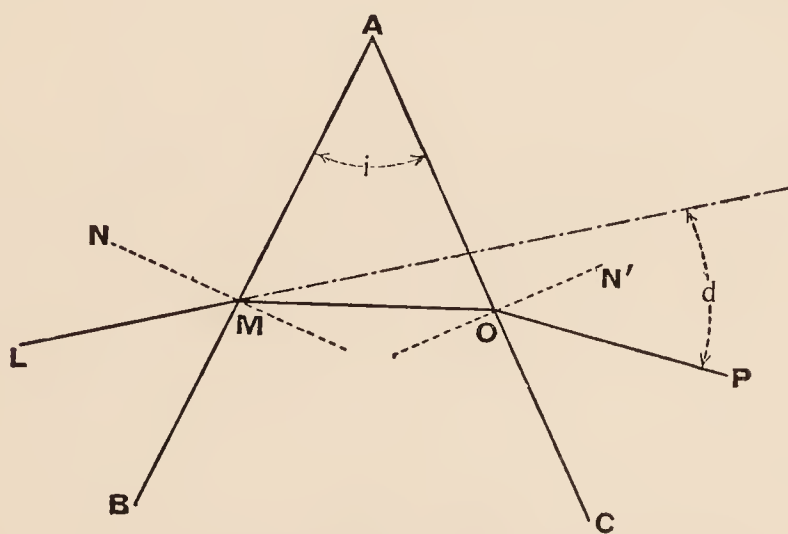


FIG. 108.

of glass. This causes a ray falling upon one face to emerge from a second face in a different direction; the change in

direction is called the *deviation* produced by the prism, and it is towards the 'base,' i. e. the third face of triangle, as may be seen from Fig. 108. Here ABC is a plan of the prism (supposed to be placed upright on one of its triangular ends), A is the refracting edge, BC the base. A ray LM falls on the first face at M, is refracted to o on the second face, and emerges in the direction OP. The angle of deviation  $d$  is the angle between LM and OP. The angle BAC is called the 'angle of the prism'—say  $\iota$ .

The course of the ray may be traced out by means of the geometry of the figure and the law of refraction, which regulates the relation between the angles of incidence and refraction at the two faces. The most important case to consider is the *symmetrical case* (shown in the figure), in which the angle of incidence LMN is equal to the angle of emergence PON', N and N' being on the normals at M and o. In this case the whole figure is symmetrical about a line bisecting the refracting angle, and the following relation may be proved to hold

$$\mu \sin \frac{\iota}{2} = \sin \frac{\iota + d}{2}.$$

This is the position in which a prism is mostly used for spectroscopic purposes, and we will therefore illustrate it by a numerical example. A prism of  $60^\circ$  angle is made from dense flint glass for which  $\mu_D = 1.6224$  (yellow)  $\mu_G = 1.6461$  (blue-violet). For these two kinds of light  $\mu \sin \frac{\iota}{2} = 0.8112$  and  $0.8230$  respectively; from this we find that  $\frac{(\iota + d)}{2} = 54^\circ.13'$  and  $55^\circ.23'$  respectively, whence  $d = 48^\circ.26'$  and  $50.46'$ . On account of the variation of refractive index with the kind of light, it is obvious that the different kinds of light will be deviated by varying amounts, i. e. will be dispersed; and the dispersion between the two colours chosen in the example is shown to be  $2^\circ.20'$ ; the whole spectrum from extreme red to extreme violet is about twice as long. (It should be noted that when the prism is in the symmetrical position for one ray it is not so for the other.)

It will be found that if the prism is not in the symmetrical position, but the angle of incidence either greater or less than that of emergence, the deviation is greater; consequently the symmetrical position is often called the *position of minimum deviation*.

When a prism is thin, i. e. has a small refracting angle, the deviation produced by it is practically the same, whatever direction (within wide limits) the light takes in passing through it; and the formula given above for the deviation may be simplified into  $d = (\mu - 1) \alpha$ . This should be compared with the formula for the strength of a lens; it will be seen that the angle of the prism plays the same part as the total curvature (sum of the curvatures of the two faces) in the lens, and that the angular deviation of the prism corresponds to the change in convergence produced by the lens. The simplified formula may be regarded as accurate enough in dealing with prismatic spectacles.

On the other hand, if the angle of a prism be made too wide, light will not pass through it. This is due to the phenomenon known as *total reflection*, which will be easily followed from this example. A prism of ordinary glass ( $\mu = 1.52$  say) is made with one angle of  $90^\circ$  and the other two  $45^\circ$  each (Fig. 109), and light allowed to fall in a parallel beam normally on one of the side faces; it therefore enters the glass without change in direction, and meets the hypotenuse face at  $45^\circ$ . If it were to be refracted into the air, we should have  $\mu \sin \iota_2 = \sin \iota_1$ , where  $\iota_2 = 45^\circ$  is the angle between the ray and the normal in glass, and  $\iota_1$  is the corresponding angle in air. But  $\mu \sin \iota_2 = 1.52 \times \sin 45^\circ = 1.075$ , and no angle can have a sine greater than unity. There is, therefore, no refracted beam, but all the light is reflected, as shown by the figure, and emerges normally to the second side face. Such a prism, therefore, forms a very efficient mirror, and is often used for that purpose. Consideration of the equation expressing the law of refraction shows that total reflection never occurs when light passes from a medium of lower to one of higher refractive index, but only in the contrary case; and in particular in passing from a medium of refractive index  $\mu$  into air when  $\mu \sin \iota_2 > 1$ , i. e.  $\sin \iota_2 > \frac{1}{\mu}$ . The value of  $\iota_2$  which has a sine equal to  $\frac{1}{\mu}$  is called the critical angle.

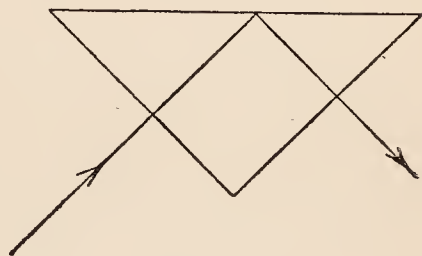


FIG. 109.

A *spectroscope* is an arrangement for separating the various



kinds of light as far as possible from one another. It consists of the following parts (Fig. 110): A slit formed by a pair of

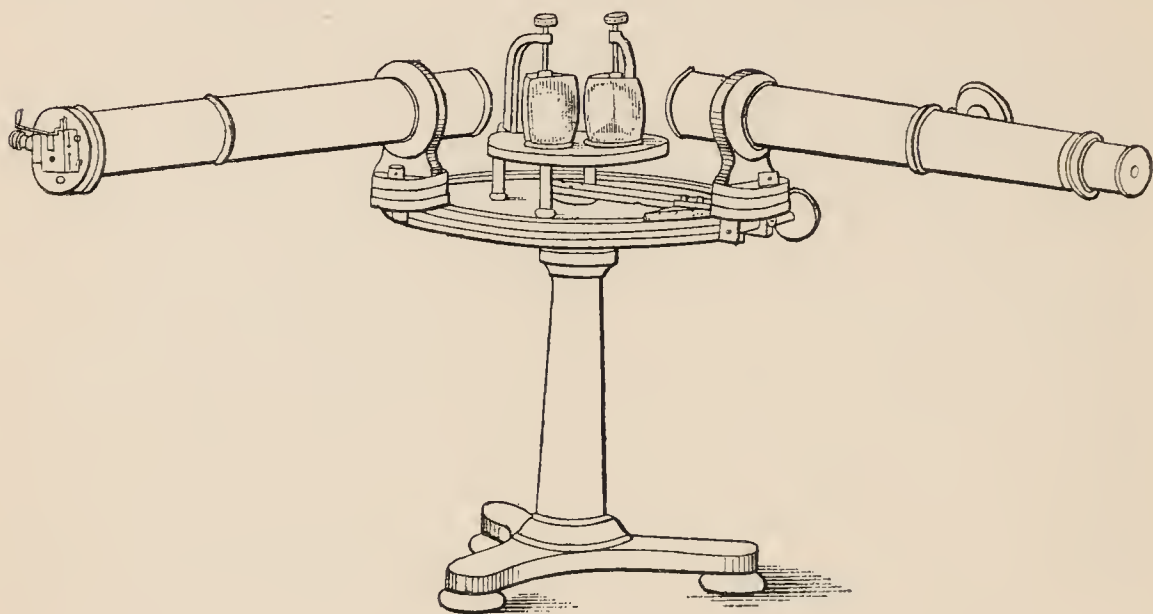


FIG. 110.

metal jaws moved by a screw to be a convenient distance apart. The length of the slit must be placed accurately vertical, and the breadth made as small as the intensity of the light employed will permit. The slit is carried by a tube called the collimator, and at the other end of this is placed the collimating lens, which is an achromatic telescope objective. The slit is moved to and fro in the tube till it is precisely in the principal focus of this lens, so that the light after traversing it falls in a parallel beam on the prism. The latter is usually carried on a rotating table, which can be accurately levelled, so that the refracting edge of the prism is vertical, i.e. parallel to the slit. (The figure shows two prisms on such a table.) Considering now one particular wave-length of light—say that from a sodium flame—since all the beam incident on the prism is parallel to itself, all that emerging from the prism is parallel to itself; it is allowed to fall on an achromatic telescope, focussed to receive parallel rays (i.e. to view distant objects). The telescope for this purpose must be capable of rotation round a vertical axis passing through the prism. If it be turned in the direction of the deviated light, a sharp yellow image of the slit will be seen (or two images side by side, if the spectroscope is good enough

to separate the two sodium rays  $D_1$  and  $D_2$ ). Now by turning the prism to and fro a position may be found for it in which the deviated image approaches nearer to the line of the collimator than in any other. This is accordingly the position of minimum deviation, or symmetrical position; if the telescope be provided with a divided circle by which to measure the angle of deviation, the refractive index of the prism can be calculated by means of the formula on p. 292. The angle of the prism can be measured on the same instrument; and for liquids, a hollow glass prism is made, closed by plane glass sides; if this be filled up with the liquid in question, it can be used like the glass prism, and the refractive index of the liquid determined.

If instead of light of one or two definite kinds, such as is given by an incandescent gas, a source of white light is used, containing radiation of all possible wave-lengths, there will be an infinite number of images of the slit side by side—i.e. actually a continuous horizontal band of light, of vertical height equal to that of the slit. It is this coloured band which, when derived from sunlight, is interrupted by narrow vertical dark spaces—the Fraunhofer lines. Of course the images of the slit formed by light of neighbouring wave-lengths overlap; the arrangement of apparatus described is designed to make the overlapping as small as possible—to obtain a so-called ‘pure’ spectrum.

The instrument in the above form is known as the spectrometer; in a spectroscope for examining the character of light from various sources, observing absorption bands, &c., it is not necessary that the prism should be placed on a rotating table, nor that the telescope should be capable of more than a small angular motion so as to pass from one end of the spectrum to the other. On the other hand, a single prism gives a short spectrum: more dispersion may be obtained by using two or more prisms in succession. Again, for minute examination of the spectrum, photography is better than a subjective method, so that a spectroscope may conveniently consist of a collimator as described, a chain of prisms permanently set up, so as to be in the position of least deviation for some standard kind of light, say the D ray, and a photographic camera arranged so as to receive the light from the prisms and throw it on a horizontal strip of sensitized film.

To compare the spectra of two lights, one is placed directly in front of the collimator, the other at the side in such a position that a totally reflecting prism can throw the beam from it on to the slit, this prism arranged so as to cover only, say, the upper half of the slit. Consequently, the upper half of the field of view of the telescope or camera receives only light from one of the sources, the lower only from the other. If one of the spectra is formed by sunlight, the Fraunhofer lines serve as a convenient means of comparison by which to determine the position of any lines seen in the other spectrum.

An instrument, such as that described above, cannot be easily moved ; for that reason portable hand spectroscopes have been devised, in which all the parts are enclosed in one tube. The instrument can therefore be directed at any object like a telescope. The usual construction of the prism in such an instrument is of alternate crown and flint prisms, turned

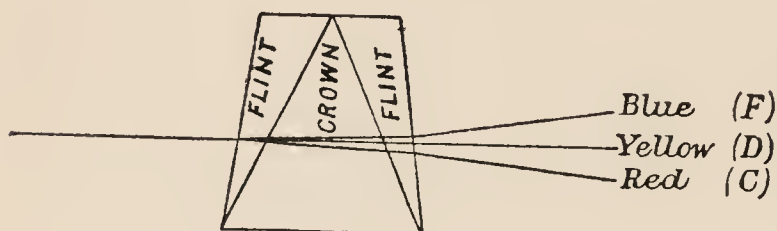


FIG. 111.

with their edges opposite ways (Fig. 111). The deviation produced by the flint may thus be neutralized by the crown, without neutralizing the dispersion. Taking the approximate formula for thin prisms,

$d = (\mu - 1) \iota$ , the theory of this may be given simply. Using, as before, suffixes  $_{CDF}$  to indicate the kind of light employed, we may take as the mean deviation  $(\mu_D - 1) \iota$ , and choose the angles of the prisms so that this quantity has the same value for the flint prism (or sum of the prisms) as for the sum of the crown prisms ; hence the D (yellow) ray will emerge from the set of prisms in the same direction as it entered. But the angular extent of spectrum (between two given rays, say C and F) produced by a single prism is the difference between the deviation of the C ray  $(\mu_C - 1) \iota$  and that of the F ray  $(\mu_F - 1) \iota$ , that is  $(\mu_F - \mu_C) \iota$ , or, as we may write it,

$$(\mu_F - \mu_C) \iota = (\mu_D - 1) \iota \div \left\{ \frac{\mu_D - 1}{\mu_F - \mu_C} \right\}$$

or, in words, the mean deviation, divided by the ratio given in the last column of the table, p. 264. As the mean deviation is to be the same for flint and crown, but the ratio is 36.9 for flint, and 60.9 for crown, the dispersion of the flint will overpower that of the crown, and a spectrum will be produced in the same direction, though not so long as that which the flint prisms alone would give.

The achromatization of objectives involves the converse problem, that



of producing deviation (change of convergence) in the light falling on them, but without dispersion, and can accordingly be treated by the same method. We have  $(\mu - 1)r$  for the strength; consequently, as before,  $(\mu_F - \mu_C)r$  for the difference in strengths for the blue and red rays; but the difference in strength of the crown lens must be neutralized by an equal difference in the opposite sense produced by the (concave) flint lens. So that  $(\mu_F - \mu_C)r$  being the same for each, their respective strengths for the mean (yellow) ray will be in proportion to the numbers given in the last column of the table, i.e. as 60.9 : 36.9. If, for example, the crown lens is of 6.09 D, and the flint of -3.69, they will achromatize and yield a compound lens of 2.4 D.

### § 7. The eye.

The optical structure of the eye is essentially as follows—

(i) The cornea, approximately spherical in shape. The radius of curvature of the anterior surface is about 8 mm., the curvature, therefore,  $1000 \div 8 = 125$  diopters. The medium that the cornea is composed of has a refractive index of 1.3376, slightly greater than that of water. The surface, it is easily seen, produces a converging effect on light incident on it (see Fig. 112), and its strength is  $125 \times (1.3376 - 1) = 42.2$  D; it accounts, therefore, for about two-thirds of the whole refracting power of the eye (p. 284): the so-called 'reduced eye' may be regarded as the cornea, somewhat strengthened so as to include the effect of the crystalline lens. Behind the cornea comes the vitreous humour, but as this has practically the same refractive index, the distinction has no optical effect.

(ii) 4 mm. behind the anterior face of the cornea comes the anterior face of the lens. The lens is 4 mm. thick, and more sharply curved behind than in front, its curvatures (in repose) being 100 D (10 mm.) in front and 167 D (6 mm.) behind. It is not, like a glass lens, made of homogeneous material, but grows denser (and more strongly refracting) towards the centre. The mean refractive index is 1.4545. According to the rule given in the footnote to p. 284, both faces of the lens produce a converging effect, which therefore augments that of the cornea.

(iii) 13 mm. behind the posterior face of the lens is the retina, the screen on which the optical image is to be formed. The space between the two is filled by the vitreous humour, which has the same refractive index as the aqueous.

The eye cannot be treated as a 'thin lens,' in which the effects of the two faces need merely be added to get the effect of the whole, for the thicknesses involved are of the same order of magnitude as the radii of curvature. For a complete optical treatment we must refer to larger treatises. The result is to show that it has a strength of about 66 D, and is so arranged that the image of a very distant object is formed on the retina, the course of the rays being that shown in Fig. 112. Here A is the

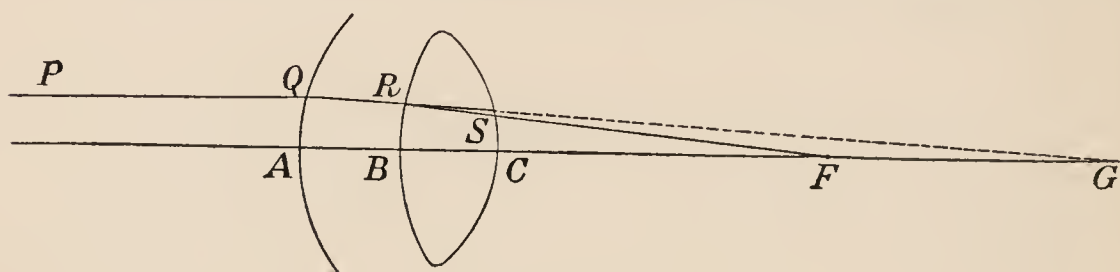


FIG. 112.

anterior surface of the cornea, BC the lens. The ray PQ is bent by the cornea into the direction QRG; hence the image of a distant object if formed by the cornea alone would fall at G. The ray, however, is bent again at R, and again at S, and reaches the axis at F, which is, consequently, the principal focus for the whole eye. It is here that the retina lies in the normal eye.

When the eye has to view an object near at hand, it is *accommodated* for the purpose by a peripheral compression of the lens, which increases its curvature, and consequently its strength. The amount of accommodation that can be effected varies with age, being greatest for children; it is usually stated at 5 diopters for adults, i. e. a distance of 20 cm.; but with effort the eye can usually see objects distinctly at shorter distances than this; this means an increase in strength of the eye from 66 to 71 D. The course of the rays in dealing with a near object may be gathered by a comparison of Fig. 102. Here the single refracting surface (representing a simplified eye) has its principal focus at  $r'$ , and forms an image at  $A'$ : in the eye the adjustment would be such that  $A'$  falls on the retina, and consequently the principal focus (which, for a distant object, is on the retina) now lies in front of it.

The chief optical defects of the eye are—

*Myopia.* The eye is too strongly curved, consequently the

image of a distant object is brought to a focus in front of the retina, and is not clearly perceived. If e.g. the eye be myopic by 2 diopters, then when in repose it is adapted to light of that divergence, that is to light from an object half a metre away; that is the 'far point of vision' for such an eye. The accommodation may be normal: in this case, instead of being able to deal with light of 5 D divergence, as the normal eye can, it would be able to deal with 7 D, i. e. with light coming from about 14 cms. distance; such an eye would, therefore, be able to see more detail in a small object than the normal eye without a magnifying glass. The remedy for myopia is to use *divergent* spectacles such as will reduce the eye to the normal strength; in the case supposed, 2 D.

*Hypermetropia.* The eye is not sufficiently curved. Hence the 'near point of vision' is abnormally far off. Thus, if an eye be hypermetropic by 2 D, and has the normal accommodation, it will at most be able to deal with  $5 - 2 = 3$  D divergence, and an object must be placed 33 cms. away to be seen. The remedy is the use of *convergent* spectacles.

*Presbyopia* is loss of accommodation, usually due to age. Its optical effect is the same as that of hypermetropia.

*Astigmatism* is unequal strength in different directions, usually due to the cornea being more curved in one meridian than another. It can be remedied by the use of astigmatic lenses, i. e. lenses made with cylindrical surfaces, instead of spherical. Such lenses have a refracting effect at right angles to the axis of the cylinder, but parallel to the axis have none, i. e. they behave like plane glass.



# INDEX

- ABBE, 277.  
Aberration, 279.  
Absolute temperature, 51, 64.  
Absorption and radiation, 271.  
— of wave, 73.  
— spectra, 277.  
Acceleration, 8.  
— of gravity, 17.  
Accommodation, 298.  
Accumulator, 198.  
Achromatism, 296.  
Aeration of blood, 175.  
Air-pump, 42.  
Alternating currents, 221.  
— dynamo, 247.  
Amalgamation of zinc, 194.  
Ammeter, 187.  
Ampère, 235.  
Ampère's rule, 233.  
Amplitude, 32.  
ANDREWS, 92.  
Anode, 189.  
Aperture of a lens, 290.  
Areometric method, 12.  
Arm, mechanics of, 28.  
Artery, flow of blood in, 119.  
Astatic galvanometer, 214.  
Atmospheric pressure, 38.  
Atomic theory, 3.  
Automatic interrupter, 243.  
Avidity, 169.  
Avogadro's law, 67, 78.  
  
Back electromotive force, 205.  
Ballistic balance, 13.  
— galvanometer, 241.  
Barlow's wheel, 237.  
Barometer, 37.  
Barometric correction of thermometer, 47.  
BECKMANN, 99.  
Beckmann thermometer, 49.  
BERGET, 70.  
BERTHELOT, 104.  
Bichromate cell, 196.  
Blood, density of, 111.  
Boiling, air bubbles in, 111.  
BOLTZMANN, 65.  
  
BOYLE, 39.  
Boyle's law, 77.  
Breaking point, 136.  
Brittleness, 136.  
BUNSEN, 105.  
Bunsen cell, 196.  
  
CAILLETET, 93.  
Calorie, 55.  
Calorimeter, 103.  
Camera, 291.  
Capacity, 250.  
— and conductance, 253.  
Capillary electrometer, 216.  
— error of barometer, 112.  
— tube, 112.  
Carbon megohm, 202.  
Cathode, 103.  
Centre of gravity, 31.  
Characteristic equation, 79.  
Charles's law, 78.  
CHAUVEAU, 61.  
Chemical equilibrium and temperature, 183.  
Chronograph, 5.  
Circulatory system, 118.  
Clark cell, 197.  
CLAUSIUS, 65.  
Clinical thermometer, 50.  
Clothing, efficiency of, 76.  
Coefficient of expansion, 88.  
COHEN, 173.  
Colloids, 128.  
Concentration, 166.  
Condensation of gases, 93.  
Condenser, 250.  
Conductance, 199.  
Conductivity, electric, 200.  
— of the soil, 71.  
— thermal, 68.  
Conductors and insulators, 185.  
Constant volume thermometer, 50.  
Contracted vein of liquid, 114.  
Convection, 72.  
Creeping of salt solutions, 112.  
Critical angle, 293.  
— point, 92.  
Cryohydrates, 143.

- Current balance, 211.  
 Curvature, 283.  
 DALTON, 65.  
 Dalton's law, 95.  
 Damping of galvanometer, 242.  
 Daniell cell, 192, 197.  
 D'Arsonval galvanometer, 213, 239.  
 Declination, 230.  
 Degrees of freedom, 171.  
 Density, 9.  
   — of solids, 12.  
 Dew point, 97.  
 Dielectric, 250.  
   — constant, 256.  
 Diffraction, 265.  
 Diffusion of dissolved substances, 122.  
 Diffusivity, 121.  
 Dilatometer, 88.  
 Diopter, 283.  
 Dip, 230.  
 Direct vision spectroscope, 296.  
 Dispersion, 263, 292.  
 Dispersive power, 296.  
 Dissipation of energy, 64.  
 Dissociation of water, 169.  
 Distribution of energy in spectrum, 272.  
 Drying agents, 175.  
 Dulong and Petit's law, 141.  
 Dynamical equivalent of heat, 55.  
 Dynamic vapour pressure measurement, 87.  
 Dynamo, 241.  
 Dyne, 16.  
 Ear, perception of sound by, 161.  
 Earth coil, 240.  
 Effective force, 15.  
 Efficiency of an engine, 61, 64.  
   — of human body, 180, 183.  
   — of lights, 272.  
 Elasticity, 80.  
   — of a liquid, 85.  
   — of solids, 129.  
 Elastic limit, 136.  
   — after-effect, 138.  
 Electrical equivalent of heat, 224.  
   — power, 193.  
 Electric heating, 223.  
 Electrolysis, 189.  
 Electrolyte resistance, 221.  
 Electrolytic dissociation theory, 99, 108, 168.  
 Electromagnet, 233, 236.  
 Electro-magnetic radiation, 258, 268.  
 Electrometer, 216.  
 Electromotive force, distribution of, 204.  
   — — of cell, 193.  
 Electromotor, 238.  
 Electrophorus, 253.  
 Electroscope, 254.  
 Electrostatic induction, 253.  
   — stress, 249.  
 Emissivity, 74.  
 Endosmotic equivalent, 125.  
 Endothermic reaction, 179.  
 Energetics, 3.  
 Energy, 1.  
   — units, 22.  
 Equilibrium, 15.  
 Erg, 20.  
 Ether, 249.  
 Evaporation from the skin, 102.  
 Expansion of solids, 138.  
 Expired air, composition of, 96.  
 Eye, 297.  
 Eye-piece, 288.  
 FARADAY, 190.  
 FICK, 17, 40.  
 Filter pump, 117.  
 Fleuss pump, 43.  
 Flexure, 134.  
 Flow of liquids, 131.  
   — of solids, 137.  
 Fluid pressure, 36.  
 Fluorescence, 273.  
 Focal length, 281, 284.  
 Food stuffs, thermal value of, 179.  
 Force, 14.  
 Fraunhofer lines, 276, 295.  
 Free energy, 183.  
   — expansion of gases, 108.  
 Frictional electricity, 253.  
 Galvanometer, 188.  
 Gas constant, 79.  
 Gases, spectra of, 273.  
   — weight of, 96.  
 GIBBS, 173.  
 Glauber salt, transition of, 174.  
 Gradient of concentration, 120.

- Grove cell, 196.  
 Hardness, 137.  
 Harmonics, 149.  
 Heart, work of the, 25.  
 Heat in electric circuit, 216.  
 Heating effect of an electric current, 186.  
 Heat of reaction, 177.  
 HELIUM, 276.  
 HELMHOLTZ, 162, 244.  
 HERIZ, 237.  
 HICKS, 13.  
 HIRN, 2, 59.  
 Homogeneous system, 164.  
 Hooke's law, 81, 130.  
 Horse-power, 24.  
 Human body, efficiency of, 61.  
 HÜRTHLE, 41.  
 Hydrates, equilibrium of, 174.  
 Hydriodic acid, equilibrium of, 167.  
 Hydrogen and oxygen, combination of, 167.  
 Hydrometer, 11.  
 Hygrometry, 97.  
 Hysteresis, 231.  
 Image, 279.  
 Incandescence, 272.  
 Index notation, 4.  
 Indicator diagram, 24.  
 Induction coil, 243.  
 Inertia, 16.  
 Infra red, 261.  
 Insulation resistance, 201.  
 Intensity of magnetization, 229.  
 — of sound, 162.  
 Internal energy, 176.  
 — energy of fluids, 101.  
 — latent heat, 107.  
 Ion, 168, 188.  
 Isothermal, 79.  
 — of a liquid, 84.  
 Isotonic method, 17.  
 JULY, 105.  
 Joule, 22.  
 JOULE, 2, 56, 108.  
 KELVIN, 65, 71, 108, 211.  
 Kinetic energy, 21.  
 KIRCHHOFF, 271.  
 Kymograph, 40.  
 LAPLACE, 153.  
 Latent heat, 102.  
 — — of fusion, 141.  
 — — of transition, 173.  
 Law of inverse squares, 251.  
 Laws of solution, 127.  
 LEES, 69.  
 Length, unit of, 3.  
 Lenz's law, 240.  
 Lever, 30.  
 Light and sound waves, 264.  
 LINDE, 94, 108.  
 Lines of force, 251.  
 — — magnetic force, 227.  
 Line spectra, 274.  
 LIPPMANN, 216.  
 Liquefaction of air, 95.  
 Liquid mixtures, vapour pressure of, 98.  
 LISTING, 282.  
 Litre, 4.  
 Longitudinal and transverse vibrations, 150.  
 Luminescence, 272.  
 Magnetic effect of current, 187.  
 — field, 229.  
 — induction, 240.  
 — moment, 228.  
 Magnetism of iron, 173.  
 Magnetization of iron, 231.  
 Magnetizing coil, 235.  
 Magneto, 241.  
 Magnetometer, 231.  
 Magnification of image, 282.  
 Magnifying glass, 287.  
 Manganin, 201.  
 Mass, standards of, 9.  
 Maximum and minimum, 180.  
 — density of water, 72.  
 MAXWELL, 65.  
 Mean free path, 66.  
 Melting point, 139.  
 — — and pressure, 140.  
 Mercury pump, 44.  
 Metabolism, 181.  
 Metastable equilibrium, 181.  
 Mho, 199.  
 MICHELSON, 3.  
 Microscope, 290.  
 Minimum deviation, 292.  
 — potential energy, 180.  
 Mirror galvanometer, 213.



Mixtures, melting point of, 142.  
 Molecular volume, 79.  
 Molecule, 65.  
 Moment of a force, 29.  
 Momentum, 12.  
 Musical scale, 148.  
  
 Neef's hammer, 245.  
 Nerve-muscle electromotive force, 210.  
 NEWTON, 270.  
 Normal salt solution, conductivity of, 201.  
 — scale of temperature, 50.  
  
 Ohm's law, 199.  
 Optical centre, 286.  
 Organ pipes, 147.  
 Oscillations, 32.  
 Oscillatory discharge, 257.  
 Osmotic pressure, 126.  
 — — at freezing point, 144.  
  
 PÉCLET, 69.  
 Pendulum, 34.  
 Perfect gas, deviation from, 83.  
 — gases, 77.  
 Permanent set, 136.  
 Perpetual motion, 62.  
 Phase, 34.  
 — of chemical system, 170.  
 Phosphorescence, 273.  
 Photometer, 262.  
 Physiological calorimeter, 61.  
 PICTET, 93.  
 Pitch of sounds, 147.  
 Plants, respiration of, 170.  
 Plasmolysis, 128.  
 Platinum thermometer, 223.  
 Polarization of electrodes, 194.  
 — of light, 268.  
 Positive and negative poles, 204.  
 Post office box, 220.  
 Potential energy, 21.  
 Potentiometer, 217.  
 Power, 24.  
 — in electric circuit, 207.  
 Pressure gradient, 115.  
 — gauge, 40.  
 — in blood vessels, 114.  
 — in branch tubes, 118.  
 — in soap bubble, 110.  
 Principal focus, 280.

Prism, 291.  
 Pulley system, 30.  
 Pulse, 119.  
  
 Quadrant electrometer, 255.  
 Quality of sounds, 148.  
 Quantity of magnetism, 228.  
  
 RAMSAY, 276.  
 RAOULT, 99.  
 Ratio of static and magnetic units, 252.  
 RAYLEIGH, 162.  
 Reaction constant, 166.  
 Reduced eye, 282.  
 Reflection and refraction of sound, 154.  
 — law of, 266, 267.  
 Refractive index, 263.  
 REGNAULT, 91, 104, 153.  
 Resistivity, 200.  
 Resonance, 157.  
 — of the ear, 163.  
 Respiration, 121.  
 Resultant of forces, 15.  
 Reversal of spectral lines, 275.  
 Reversible reaction, 165.  
 Rigidity, 131.  
 Röntgen rays, 258.  
 ROWLAND, 56.  
 RUMFORD, 2.  
  
 Salt solutions, freezing point of, 143.  
 Sap, rise of, 112, 128.  
 Saturation pressure, 85.  
 Screw, 29.  
 Second, 4.  
 Semi-permeable membrane, 125.  
 Sensitiveness of the eye, 277.  
 Series and parallel arrangement, 207.  
 Shear, 130.  
 Short circuit, 194, 208.  
 Shunt, 208.  
 Simple harmonic motion, 33.  
 Sine curve, 35.  
 Size of atoms, 66.  
 — of drops, 110.  
 Soft tissues, elasticity of, 132.  
 Solutions, vapour pressure of, 99.  
 — boiling point of, 99.  
 Specific gravity bottle, 10.  
 — heat, 55.

- Specific heat of gases, 102, 107.  
 Spectacle lenses, 299.  
 Spectroscope, 269, 294.  
 Sprengel pyknometer, 10.  
 Stability and temperature, 174.  
   — of solid and liquid, 141.  
 Stable equilibrium, 31.  
 Static vapour pressure measurement, 86.  
 Steam engine test, 60.  
 STEFAN, 270.  
 Stefan's law, 75.  
 STOKES, 270.  
 Strain, 80.  
 Sulphur, transformation of, 172.  
  
 Tangent galvanometer, 212, 235.  
 Telescope, 289.  
 Temperament, 148.  
 Temperature, effect of reaction, 169.  
   — gradient, 70.  
   — and molecular theory, 67.  
   — of visibility, 261.  
 Terrestrial magnetism, 230.  
 Theory of exchanges, 270.  
 Therm, 59.  
 Thermal capacity, 55.  
   — after effect, 139.  
   — stress on solids, 139.  
 Thermo-couple, 225.  
 Thermometer, 46.  
 Thermostat, 52.  
 Thin lens, 285.  
 Third law of motion, 13.  
 THOMSON, J. J., 66.  
 Time marker, 7.  
 Tin, transition of, 174.  
 Torque, 30.  
 Total reflection, 293.  
 Transformer, 247.  
 Transition point, 173.  
 Triple point, 172.  
 True and mean coefficient, 89.  
 True specific heat, 102.  
 Tuning-fork (electro-magnetic), 4.  
 Twist, 132.  
  
 Ultra violet, 261.  
   — — spectrum, 269.  
 Under cooled liquid, 141.  
 Unit of current, 235.  
  
 Vacuum jacket, 76.  
   — tube, 258.  
 VAN 'T HOFF, 99.  
 Van't Hoff's analogy, 127.  
 Vectors, 8, 26.  
 Velocity, 8.  
   — head, 113, 116.  
   — of light, 263.  
   — of reaction, 166.  
   — of sound, 153.  
 Ventilation, 97.  
 Vibrations in wires, 146.  
 Virtual velocities, principle of, 28.  
 Viscosity of liquids, 115, 116.  
 Visible spectrum, 261.  
 Volt, 193.  
 Voltage, 192.  
 Voltmeter, 190.  
 Voltmeter, 216.  
 Vowel-sounds, 160.  
  
 Water pump, 45.  
 Water, resistance of pure, 201.  
 Watt, 24.  
 Wave, propagation of, 152.  
   — front, 265.  
   — length, 36.  
   — length and temperature, 271.  
   — length of light, 261.  
   — length standard, 3.  
   — motion, 35.  
 Wehnelt interrupter, 259.  
 Weight thermometer, 90.  
 Wet and dry bulb thermometer, 98.  
 Wheatstone's bridge, 219.  
 Wimshurst machine, 254.  
 Work, 20.  
   — diagram, 23.  
  
 Young's modulus, 133.











